



Morphogénèse et variabilité au sein du système oculo-moteur

Céline Paeye

► To cite this version:

Céline Paeye. Morphogénèse et variabilité au sein du système oculo-moteur. Philosophie. Université Charles de Gaulle - Lille III, 2012. Français. NNT : 2012LIL30011 . tel-01124104

HAL Id: tel-01124104

<https://theses.hal.science/tel-01124104>

Submitted on 22 Apr 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Université Lille III - Lille Nord de France



Laboratoire URECA EA 1059

Ecole Doctorale Sciences de l'Homme et de la Société

Morphogenèse et variabilité au sein du système oculo-moteur

Céline PAEYE

THESE

Soutenue en Juillet 2012

En vue de l'obtention du diplôme de
doctorat en Psychologie

Sous la direction de
Pr. Jean-Claude DARCHEVILLE
& Dr. Laurent MADELAIN

Composition du Jury :

Pr. Karine DORE-MAZARS

Dr. Denis PELISSON

Dr. Laurent GOFFART

Pr. Jean-Claude DARCHEVILLE

Dr. Laurent MADELAIN



Remerciements

Tout d'abord, je remercie les membres du jury, le Professeur Karine Doré-Mazars, le Docteur Denis Pélisson et le Docteur Laurent Goffart d'avoir accepté d'évaluer mon travail de thèse, et particulièrement le Professeur Jean-Claude Darcheville pour ses conseils avisés.

Je ne remercierai jamais assez Laurent Madelain, dont j'ai sans doute mis la patience à très rude épreuve (comme il a parfois éprouvé la mienne) lors de ces sept années pendant lesquelles il m'a permis de découvrir le monde de la recherche. Le remercier pour les nombreuses connaissances qu'il m'a transmises, le soutien précieux qu'il m'a apporté et la confiance qu'il m'a accordée. Le remercier aussi pour les horizons qu'il m'a ouverts.

Grâce à lui, le marathon de la thèse a été l'une des plus belles aventures qu'il m'a été permis de vivre aujourd'hui (si, si...).

Merci à vous, Maman et Papa, de m'avoir encouragée à chaque kilomètre, prodigué du réconfort à chaque ravitaillement, aidée à passer les fameux « murs » qui ont jalonné ce parcours. Merci d'avoir été et d'être là.

Merci à toi, Marie, avec qui j'ai grandi et avec qui j'espère vivre encore de nombreux moments de complicité.

Merci aussi à toi, Joëlle, qui me connais si bien et qui sais me tendre la main quand je m'essouffle : *with your help* « *what I want and aim at seems less confoundedly difficult* ». Merci pour ton écoute et pour nos discussions interminables qui n'ont pas de prix pour moi.

Je tiens à remercier également l'équipe des doctorants d'URECA avec qui j'ai passé de bons moments. Ceux qui ont déjà remis leur dossard à l'arrivée : Pauline, Nora, Ludivine, Olivier, Marion et Jérôme. Ceux qui sont comme moi dans la dernière ligne droite : Jer, Antoine, Stef et Eva. Ceux qui ont pris leur départ un peu plus tard : Karinne, Aurélien, Xavier, Lucie, Mariama, Ewa.

Je voudrais aussi exprimer ma reconnaissance envers tous les participants de mes longues expériences, pour leur disponibilité et leur endurance.

Une autre équipe m'a permis de concrétiser mes aspirations : celle avec qui j'ai parcouru les routes et chemins du Nord, celle avec qui je me sens bien, tout simplement – Fred, Suzanne, Isa, Titi, Piou², Cleme, Jean-Luc, Caribou, les Dicks, la Gazelle, Céline (parce que tu n'aimes pas ton surnom), et les autres... Merci à Clément, David et Guillaume, car c'est un peu grâce à vous que j'ai pu prendre le départ. Merci à vous tous pour vos encouragements et pour tous les moments d'authenticité que nous avons partagés ensemble.

Résumé : Morphogenèse et variabilité au sein du système oculo-moteur

Un vaste corpus de recherches computationnelles est consacré à l'analyse du contrôle moteur. Cette approche, qui étudie l'adaptation saccadique comme un modèle d'apprentissage sensori-moteur, invoque des mécanismes spécialisés dans la calibration de gains internes ou l'optimisation de fonctions de coûts. Cette thèse propose une explication différente de l'apprentissage oculo-moteur en montrant que les distributions d'amplitudes saccadiques peuvent être modifiées par leurs conséquences.

D'abord, nous avons manipulé la médiane du gain saccadique. Nous avons éliminé l'erreur rétinienne post-saccadique et présenté des stimulus sonores ou visuels après les saccades répondant à des critères d'amplitude. Ce renforcement opérant a induit des changements de gain similaires à l'adaptation saccadique obtenue par le paradigme conventionnel de double saut.

Puis nous avons montré qu'une part de la variabilité peut être contrôlée par un apprentissage opérant. Différents niveaux de variabilité de gain ont été renforcés avec un son. Les changements obtenus suggèrent que la variabilité ne résulte pas uniquement d'un bruit interne. Enfin, dans une tâche de recherche visuelle, la présentation de la cible en guise d'agent renforçateur a aussi permis d'augmenter puis de réduire la variabilité des amplitudes saccadiques. Cela signifie que voir la cible est une conséquence qui contrôle (en partie) les propriétés des saccades.

Ces études montrent qu'un processus général d'apprentissage, qui dépend des conséquences fonctionnelles des saccades, peut guider les changements de leur amplitude. Cela pourrait s'avérer crucial pour maintenir la précision des saccades tout au long de la vie.

Mots-clés : saccades oculaires, contrôle moteur, variabilité comportementale, apprentissage opérant, façonnage, humain.

Abstract : Morphogenesis and variability in the oculomotor system

A large body of computational research has been devoted to understanding motor control. This approach, using saccadic adaptation as an example of sensorimotor learning, postulates specialized mechanisms responsible for a motor calibration or optimization processes of cost functions. This thesis dissertation provides an alternative view of oculomotor learning by showing that saccadic amplitude distributions can be altered by their own consequences.

First we manipulated saccadic gain medians. We eliminated the postsaccadic retinal error (either by extinguishing the target during the saccade or by stabilizing its image on the fovea) and provided auditory or visual consequences after the saccades meeting amplitude criteria. This reinforcement procedure could induce gain changes similar to saccadic adaptation obtained via a conventional double-step paradigm.

Then we showed that part of saccadic amplitude variability may be controlled by operant learning. Saccadic gain distributions were reinforced with a tone depending on variability criteria. Our data suggested that oculomotor variability may not result solely from an internal random noise. Finally we designed a new paradigm involving a visual search task in which presenting a target was effective to reinforce various saccadic amplitude variability levels. This means that in real life, seeing the target is a consequence controlling (at least in part) saccadic properties.

These results show that a general operant learning process, depending on the functional consequences of eye movements, can guide changes in saccadic amplitude. This may be critical to maintain saccadic accuracy during the lifespan.

Key words : ocular saccades, motor control, behavioral variability, operant learning, shaping, human.

Laboratoire URECA - EA 1059

Unité de REcherche en sciences Cognitives et Affectives



Université Charles-de-Gaulle Lille III
Domaine universitaire du Pont de Bois
BP 14959653 Villeneuve d'Ascq Cedex
France

Tables des matières

INTRODUCTION	10
Vue d'ensemble de la thèse	12
CONCEPTS & CADRE THEORIQUE	14
I. L'apprentissage oculo-moteur selon les théories computationnelles : l'exemple de l'adaptation saccadique	15
1.1. L'adaptation saccadique, un exemple d'apprentissage sensori-moteur	15
1.2. La nature du signal d'erreur à l'origine de l'adaptation saccadique	19
1.3. Deux modèles internes <i>adaptatifs</i> pour rendre compte de l'adaptation saccadique	28
1.4. Le système oculo-moteur, un système adaptatif qui « apprendrait à effectuer des mouvements optimaux »	35
1.5. Quelques limites des théories computationnelles du contrôle moteur	39
II. L'apprentissage oculo-moteur selon l'analyse expérimentale du comportement : un mécanisme général d'apprentissage	42
2.1. Les principes fondamentaux de l'apprentissage opérant	42
2.2. La « sélection de la cible » sous l'angle de la théorie du renforcement opérant	44
2.3. Les saccades oculaires possèdent des dimensions opérantes	47
2.3. La variabilité comportementale, une dimension opérante	56
2.4. Hypothèses des recherches : le renforcement opérant, un mécanisme général d'apprentissage permettant de modifier la distribution des amplitudes des saccades oculaires	64
ETUDE 1 : MODIFICATION DU GAIN SACCADIQUE PAR LES CONTINGENCES DE RENFORCEMENT	67
I. Méthode	69
1.1. Participants	69
1.2. Equipement et stimulus	69
1.3. Le façonnage par procédure de renforcement percentile ciblé	70
1.4. Procédures des différentes expériences	73
1.5. Acquisition et analyse des données	78
II. Résultats	81
2.1. Expérience 1 : façonnage d'une diminution du gain saccadique avec agent renforçateur auditif	81
2.2. Expérience 2 : façonnage d'une augmentation du gain saccadique avec agent renforçateur auditif	83
2.3. Expérience 3 : façonnage d'une diminution du gain saccadique avec agent renforçateur visuel	86
2.4. Expérience 4 : façonnage d'une diminution du gain saccadique avec agent renforçateur auditif et stabilisation de l'image de la cible à l'issue de chaque essai	88

2.5. Expérience 5 : comparaison entre les adaptations saccadiques induites par la procédure de renforcement et par le paradigme conventionnel	90
2.6. Remarques générales sur les résultats des expériences de façonnage.....	96
III. Discussion.....	98
3.1. Les contingences de renforcement sont responsables des changements de gain saccadique.....	99
3.2. Différences entre les changements de gain résultant des procédures de façonnage et ceux produits par le paradigme de double saut.....	101
3.3. Similarités entre les changements de gain résultant des deux types de procédures	106
IV. Conclusion de la première étude : le contrôle du gain saccadique par les contingences de renforcement	113
ETUDES 2 & 3 : RENFORCEMENT DE LA VARIABILITE DE L'AMPLITUDE SACCADIQUE.....	114
I. Méthode de l'étude 2 : renforcement de la variabilité du gain de saccades horizontales.....	117
1.1. Participants	117
1.2. Equipement et stimulus	117
1.3. Procédures des expériences.....	117
1.4. Acquisition et analyse des données	122
II. Résultats de l'étude 2.....	124
2.1. Expérience 1 : étendue du contrôle de la variabilité opéré par les contingences de renforcement.....	124
2.2. Expérience 2 : la contingence entre les mouvements oculaires et leurs conséquences est-elle nécessaire pour contrôler la variabilité ? <i>Yoke control</i>	126
2.3. Analyse des <i>U values</i>	128
III. Discussion de l'étude 2.....	131
3.1. L'influence éventuelle de la fréquence de renforcement.....	132
3.2. Analyse des latences saccadiques.....	135
3.3. Analyse de la <i>main sequence</i>	136
IV. Méthode de l'étude 3 : renforcement de la variabilité de l'amplitude des saccades dans une tâche de recherche visuelle	139
4.1. Objectifs des deux expériences et caractéristiques des participants	139
4.2. Equipement et stimulus	140
4.3. Procédures des expériences.....	142
4.4. Acquisition et analyse des données	146
V. Résultats de l'étude 3.....	148
5.1. Expérience 1 : la nécessité de la contingence entre les mouvements oculaires et leurs conséquences pour contrôler la variabilité – <i>Yoke control</i>	148
5.2. Expérience 2 : test du transfert de l'apprentissage à des situations de recherche visuelle conventionnelle	151
VI. Discussion de l'étude 3.....	156
6.1. Mise en place de contrôles discriminatifs de la variabilité	157
6.2. Une influence indirecte des taux de renforcement ?	160

6.3. Les effets de la procédure de renforcement opérant sur les durées de fixation	163
6.4. Analyse de la <i>main sequence</i>	166
VII. Conclusion des études 2 & 3 : le contrôle de la variabilité de l'amplitude saccadique par les contingences de renforcement.....	168
DISCUSSION & PERSPECTIVES	170
I. La morphogenèse au sein du système oculo-moteur, un processus de sélection comportementale	173
1.1. De l'origine des variations comportementales.....	173
1.2. Renforcement direct de la variabilité opérante & façonnage : prolongement des recherches actuelles.....	175
1.3. Du développement du contrôle des mouvements oculaires	177
II. Intégration de la « valeur » des conséquences des saccades dans les modèles computationnels du contrôle moteur	180
2.1. L'effet de la manipulation des conséquences des saccades sur les différentes propriétés de ces mouvements : un consensus ?	180
2.2. La « valeur » des conséquences saccadiques : un rôle central, des définitions différentes selon les approches théoriques.....	182
III. Conclusion	190
REFERENCES.....	191
ANNEXES	202
Annexe 1 : Formulaire de consentement.....	202
Annexe 2 : Consignes de l'étude 1	203
Annexe 3 : Consignes de l'étude 2.....	204
Annexe 4 : Consignes de l'étude 3.....	205
Annexe 5 : Articles des études 1 et 2.....	206

INTRODUCTION

Le voyageur qui repère quel est son terminal d'embarquement sur le panneau d'affichage d'un aéroport, l'amateur d'art qui compare un tableau original à sa copie ou bien le radiologue qui détecte une anomalie sur un cliché ont ceci en commun qu'ils émettent tous des comportements visuellement guidés, c'est-à-dire des actions contrôlées par un signal visuel. Détecter rapidement des changements dans l'environnement peut même s'avérer crucial, par exemple pour l'animal menacé par un prédateur ou pour l'automobiliste qui doit éviter des dangers.

Dans son étude comparative, Land (1999) observe que même s'ils sont éloignés d'un point de vue phylogénétique, les vertébrés dotés d'un système visuel coordonnent les mouvements de leur corps, de leur tête et/ou de leurs yeux de telle sorte que leur regard reste stabilisé sur la cible. Les mouvements réalisés pour acquérir l'information visuelle dépendent étroitement des caractéristiques physiologiques du système visuel, et plus particulièrement de la fovéa – la partie centrale de la rétine où la densité des photorécepteurs est la plus élevée et où l'acuité est maximale.

Chez les primates, on distingue plusieurs types de mouvements des yeux (Dodge, 1903; Krauzlis, 2008). D'une part, des réflexes oculaires compensent les déplacements du corps, de la tête et/ou de l'environnement afin de limiter la dégradation de l'acuité visuelle liée au glissement de l'image sur la rétine. D'autre part, les primates émettent des mouvements oculaires volontaires, alternant des périodes de fixation et des changements d'orientation du regard, afin d'inspecter leur environnement.

Plusieurs types de mouvements volontaires permettent de maintenir ou d'amener l'image des objets sur la fovéa (qui mesure environ 2 mm de diamètre chez l'être humain) et ainsi d'obtenir une vision précise de régions limitées de l'espace visuel. Ainsi, lorsqu'une cible visuelle se déplace lentement dans un environnement stable, des mouvements continus, dit de poursuite oculaire, sont émis. La trajectoire de la poursuite est déterminée par le déplacement de la cible, l'image de l'objet est alors continuellement projetée sur la fovéa, ce qui minimise le flou qui pourrait résulter du glissement de l'image sur la rétine. Quant aux mouvements de vergence, ils correspondent aux mouvements disjonctifs pendant lesquels les yeux se déplacent dans des directions opposées, quand le regard se porte vers des objets placés à différentes distances. Enfin, les saccades oculaires, sur lesquelles porte cette thèse, sont des mouvements conjugués des yeux qui orientent

très rapidement la fovéa sur l'image des cibles visuelles – chez l'adulte humain, une saccade de 10° dure environ 40 à 50 ms et la vitesse saccadique peut atteindre 500°/s (Becker, 1989).

Les saccades oculaires présentent l'avantage de faire l'objet d'une littérature très abondante : les études comportementales, les investigations neurobiologiques ou encore les modélisations computationnelles entre autres, tirent profit des connaissances accumulées et de la facilité grandissante avec laquelle il est possible de mesurer avec précision les propriétés de ces mouvements.

De plus, le fait que les yeux ne soient commandés que par trois paires de muscles rend les comportements oculo-moteurs relativement simples à étudier comparativement aux autres comportements sensori-moteurs. Étant donné l'étendue très limitée de la fovéa, les saccades oculaires requièrent un contrôle précis afin que les images des objets puissent s'y projeter. Par conséquent, l'étude du système oculo-moteur, et des saccades oculaires en particulier, constitue un moyen privilégié de sonder les mécanismes neuronaux du contrôle moteur ainsi que d'autres processus complexes tels que les interactions sensori-motrices ou encore des processus cognitifs comme l'attention, la perception ou la prise de décision (ex. : Glimcher, 2003; Körding, Tenenbaum, & Shadmehr, 2007; Krauzlis, 2008; Leigh & Kennard, 2004; Wolpert & Ghahramani, 2000).

Par ailleurs, le système saccadique possède des propriétés adaptatives qui lui permettent de maintenir l'acuité visuelle quand des changements surviennent dans l'environnement ou dans l'organisme : on parle alors en effet d'adaptation saccadique. En manipulant expérimentalement cet environnement, il est possible d'étudier la manière dont le système oculo-moteur modifie ses mouvements en conséquence, autrement dit, il est possible d'étudier une forme d'apprentissage sensori-moteur (ex. : Collins & Doré-Mazars, 2009; Hopp & Fuchs, 2004; Pélisson, Alahyane, Panouillères, & Tilikete, 2010).

L'étude des saccades et de leur adaptation se révèle donc pertinente pour plusieurs raisons. Méthodologiques d'une part, étant donné la facilité d'accès à ce comportement et la possibilité d'en observer les modifications en laboratoire. Théoriques d'autre part, dans la mesure où l'adaptation saccadique constitue un bon modèle d'apprentissage sensori-moteur.

Vue d'ensemble de la thèse

La première partie de cette thèse contraste deux conceptions théoriques différentes du contrôle et de l'apprentissage oculo-moteurs.

En premier lieu l'approche computationnelle, et en particulier certaines théories du contrôle optimal, seront abordées au travers du phénomène de l'adaptation saccadique. Celle-ci est classiquement induite en laboratoire via le paradigme de *double saut*, élaboré par McLaughlin en 1967, qui consiste à faire disparaître la cible pendant que l'œil effectue une saccade dans sa direction et à la faire réapparaître après un saut de quelques degrés avant la fin de la saccade. Au fil des essais, l'amplitude des saccades est modifiée de telle sorte que la fovéa atteigne immédiatement l'image de la cible à l'issue du saut intra-saccadique, permettant ainsi que la vision de cette cible soit nette.

Si ce paradigme d'adaptation saccadique a donné lieu à de nombreuses publications dans le domaine de la psychologie cognitive, ce n'est pas le cas dans celui de l'analyse expérimentale du comportement. Cette dernière a d'ailleurs peu étudié l'apprentissage sensori-moteur, ce qui peut paraître surprenant dans la mesure où son objet concerne explicitement la modification des comportements – plus précisément, l'apprentissage opérant.

L'analyse expérimentale du comportement peut apporter un éclairage différent sur l'adaptabilité du système oculo-moteur dans la mesure où elle centre ses investigations sur le contrôle de l'apparition, du maintien et de la sélection des comportements par leurs propres conséquences. Selon cette théorie, c'est parce que dans le passé certaines réponses ont été suivies de conséquences bénéfiques pour l'organisme que leur probabilité d'apparition augmente. On dit alors qu'un comportement a été renforcé, qu'il est contrôlé par des contingences de renforcement (c'est-à-dire l'ensemble des relations entre les comportements, leurs antécédents et leurs conséquences). Dès lors, la morphogenèse, l'évolution des formes comportementales, est étudiée en tant que changements comportementaux résultant des modifications de ces contingences (Baer & Rosales-Ruiz, 2003; Skinner, 1953).

En second lieu seront donc présentés les éléments théoriques permettant d'envisager le contrôle et l'apprentissage oculo-moteurs sous l'angle de l'analyse expérimentale du comportement. Cette démarche amène à examiner la question de la variabilité, qui constitue un point de divergence entre les deux approches. Nous verrons que si les théories du contrôle optimal la considèrent comme résultant d'un bruit interne, comme un élément que le système chercherait à minimiser, ce n'est pas le cas pour l'analyse expérimentale du comportement. Pour elle, la variabilité peut être contrôlée (du moins en partie), voire renforcée par l'environnement.

Les parties suivantes de cette thèse présentent des données expérimentales en faveur du contrôle et de la modification de l'amplitude des saccades oculaires par les contingences de renforcement. Ces contingences peuvent être manipulées parce que les mesures effectuées en temps réel permettent d'éliminer les conséquences visuelles habituelles des saccades (par exemple en éteignant la cible au début du mouvement). L'apparition de stimulus à l'issue de certaines saccades, comme des sons indiquant que le mouvement est correct, peut alors servir d'agent renforçateur lors de l'application de procédures opérantes.

La seconde partie présente une étude testant l'hypothèse selon laquelle la mise en place de procédures de renforcement opérant permet de modifier le gain des saccades (étude 1). Dans la mesure où ces changements sont comparables à ceux qui sont induits par le paradigme conventionnel de double saut, cette recherche va à l'encontre des théories computationnelles actuelles qui postulent l'existence de mécanismes spécialisés, responsables de la calibration d'un gain interne. Nos données plaident plutôt en faveur de l'implication d'un processus général d'apprentissage – un apprentissage opérant.

Ce processus général repose sur la sélection : les saccades qui sont suivies de conséquences appropriées pour l'organisme (comme la vision nette de la cible) ont une plus forte probabilité d'apparition que celles qui ne sont pas suivies de telles conséquences. Tout comme la sélection naturelle, la sélection comportementale s'appuie sur l'existence de variations – qui ne sont donc pas forcément délétères (Donahoe, 2003; Donahoe, Burgos, & Palmer, 1993; Skinner, 1981).

La troisième partie de cette thèse est consacrée à deux autres études qui ont pour objectif de montrer que la variabilité de l'amplitude des saccades peut dépendre en partie d'un apprentissage opérant. Ainsi, l'application de procédures de renforcement permet d'induire expérimentalement différents niveaux de variabilité : d'une part en ce qui concerne le gain de saccades horizontales (étude 2), d'autre part, en ce qui concerne l'amplitude de saccades réalisées dans des situations de recherche visuelle pendant laquelle l'observateur doit trouver une cible (étude 3).

Puisque les manipulations des conséquences des mouvements oculaires se sont avérées efficaces pour modifier les propriétés des saccades, ces études attestent le rôle fonctionnel de ces conséquences dans l'apprentissage oculo-moteur. Elles suggèrent donc que cet apprentissage relève des mêmes principes que ceux qui gouvernent les autres comportements opérants.

CONCEPTS &
CADRE THEORIQUE

I. L'apprentissage oculo-moteur selon les théories computationnelles : l'exemple de l'adaptation saccadique

L'adaptation saccadique désigne une forme d'apprentissage sensori-moteur permettant au système oculaire de maintenir l'adéquation entre l'amplitude des saccades et l'excentricité de la cible.

L'amplitude des saccades correspond à la distance parcourue par l'œil entre deux points de fixation. Les saccades oculaires permettent à l'œil d'amener la fovéa sur l'image de l'objet sélectionné. Le gain saccadique, qui désigne le rapport entre l'amplitude de la saccade et l'excentricité de la cible, est alors proche de 1. Si, à plusieurs reprises, les saccades ne permettent pas de centrer l'image de la cible sur la fovéa, le gain saccadique s'en trouvera progressivement modifié.

L'adaptation saccadique s'observe suite à la modification du système oculo-moteur lui-même ; elle peut également être induite en laboratoire en quelques minutes grâce au paradigme expérimental de double saut. Après la description de ce paradigme, sera abordée la manière dont les théories computationnelles classiques rendent compte de cette adaptation.

Ces approches se sont également appuyées sur les régularités des paramètres des saccades (en dehors de ce paradigme de double saut) pour formaliser leur théorie du contrôle et de l'apprentissage oculo-moteur, ce que nous évoquerons à la fin de cette première partie.

1.1. L'adaptation saccadique, un exemple d'apprentissage sensori-moteur

1.1.1. L'adaptation saccadique consécutive aux changements des muscles extra-oculaires

Outre la croissance ou le vieillissement, des pathologies oculaires comme des paralysies des muscles oculo-moteurs, peuvent rendre les saccades dysmétriques (c'est-à-dire qu'il existe alors un écart entre l'excentricité des cibles et l'amplitude des saccades).

Par exemple, suite à une paralysie affectant son œil gauche, le patient de l'étude de Kommerell, Olivier et Theopold (1976) émettait des saccades « orthométriques » (c'est-à-dire

permettant que la fovéa atteigne l'image de la cible) avec cet œil atteint et des saccades hypermétriques avec l'œil droit (menant la fovéa plus loin que l'image de la cible, induisant un gain supérieur à 1). Après avoir couvert l'œil gauche pendant trois jours, le gain des saccades de l'œil droit est redevenu normal. Ce résultat est corroboré par l'étude d'Abel, Schmidt, Dell'Osso, et Daroff (1978) dont le patient souffrait d'une paralysie du nerf moteur oculaire droit. L'œil droit, touché par la paralysie, effectuait des saccades hypométriques (dont les amplitudes ne sont pas assez grandes pour que la fovéa atteigne l'image de la cible, le gain étant alors inférieur à 1). Forcer le patient à utiliser cet œil en couvrant l'œil sain (l'œil gauche) a induit une augmentation du gain des saccades, et ce pour les deux yeux, en un peu plus d'un jour et demi. Puis, c'est l'œil droit affecté par la paralysie qui était couvert, ce qui avait pour conséquence de réduire l'amplitude des saccades de l'œil sain, restaurant le gain initial.

Ces changements de gain observés chez l'homme suite à des lésions naturelles ont été induits chez le singe par des lésions expérimentales des muscles extra-oculaires (Scudder, Batourina, & Tunder, 1998). En utilisant cette même méthode consistant à couvrir l'œil intact puis l'œil lésé pendant plusieurs jours, des changements progressifs (une augmentation puis une diminution) du gain saccadique ont pu être obtenus.

Les auteurs de ces recherches sur les changements du système oculo-moteur concluent en termes d'adaptation : en modifiant l'innervation des muscles oculaires, le système parvient à compenser les faiblesses musculaires, préservant ainsi la précision des saccades. Par ailleurs, les résultats de Scudder et al. (1998) suggèrent aussi que cette fonction adaptative est semblable chez les hommes et les singes.

1.1.2. L'adaptation saccadique induite de manière non invasive

L'adaptation saccadique peut être observée en laboratoire chez le sujet sain grâce au paradigme de double saut, introduit par McLaughlin en 1967.

Pendant que le sujet effectue des saccades vers une cible visuelle, cette cible est déplacée (voir figure 1). Au fil des essais, l'amplitude des saccades est modifiée de telle sorte que l'œil s'approche de plus en plus de la cible à l'issue de son second saut. Si la direction du saut intra-saccadique de la cible est opposée à celle de la saccade principale, le gain (calculé par rapport à la taille du premier déplacement de la cible) diminue progressivement. En revanche, si la cible est déplacée dans le même sens que la saccade principale, le gain augmente. Puis, lors d'une phase de

désadaptation, c'est-à-dire lorsque la cible cesse de faire un saut intra-saccadique, le gain revient à son niveau initial.

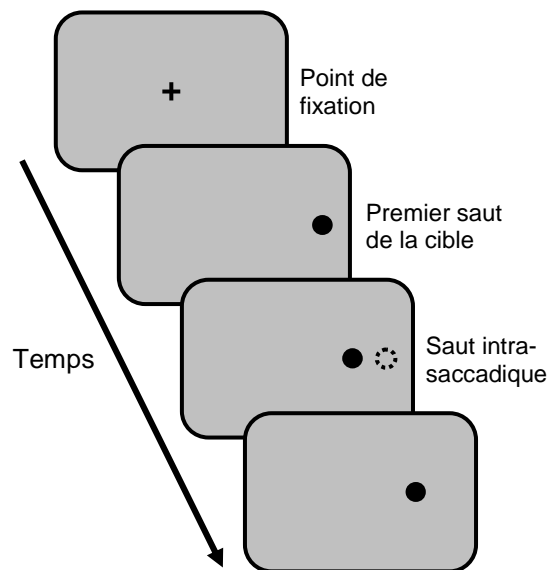


Figure 1. Déroulement d'un essai du paradigme conventionnel de double saut, avec un saut intra-saccadique dans la direction opposée à la celle du premier saut de la cible (donc visant à réduire le gain des saccades).

Ce paradigme de double saut a été maintes fois répliqué, chez l'être humain comme chez le singe : les effets en sont donc bien établis (voir Collins & Doré-Mazars, 2009; Hopp & Fuchs, 2004; Pélisson et al., 2010, pour des revues de littérature). La figure 2 illustre les décours temporels de l'adaptation et de la désadaptation observées chez un singe, ainsi que les positions de la cible et celle de l'œil à différents moments de l'expérience.

Cette procédure présente l'avantage de pouvoir induire l'adaptation saccadique de manière non invasive, mais aussi de manière plus rapide que ce qui est observé lors de dysfonctionnements des muscles extra-oculaires. Ainsi, McLaughlin (1967) ou encore Miller, Anstis et Templeton (1981) par exemple obtiennent chez l'humain une diminution du gain en une dizaine d'essais seulement, lorsque la cible répète toujours les mêmes sauts initial et intra-saccadique.

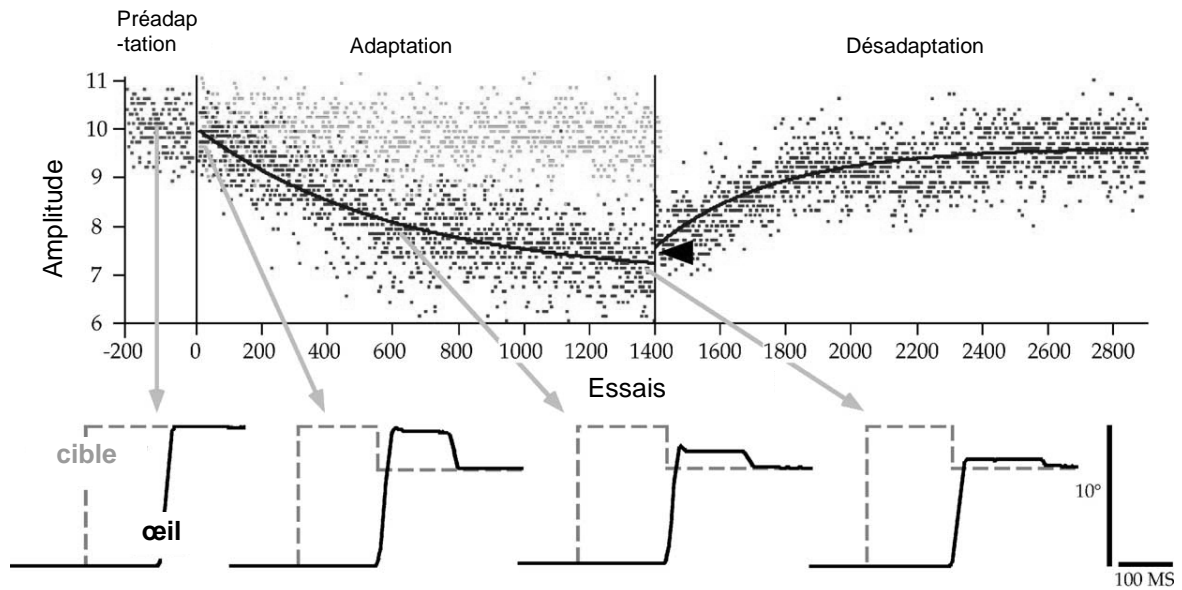


Figure 2. Déroulement temporel de l'adaptation saccadique et de la désadaptation (points gris foncé) chez le singe, lorsque la cible effectue un premier déplacement horizontal de 10° puis un saut intra-saccadique dans la direction opposée. Les points gris clair représentent les essais réalisés dans une direction où la cible n'effectue aucun saut intra-saccadique. Les données de l'adaptation et de la désadaptation sont résumées par une fonction exponentielle. En dessous de ce graphique sont représentés des exemples de positions de l'œil par rapport à celles de la cible à différents moments de l'expérience. (source : Hopp & Fuchs, 2004.)

Néanmoins, plusieurs études révèlent que si les déplacements initiaux de la cible varient, l'adaptation saccadique est considérablement ralentie (Miller et al., 1981; Albano & King, 1989). Dans leur recherche réalisée chez des singes, Scudder et al. (1998) comparent expérimentalement l'adaptation saccadique obtenue par le paradigme de double saut et celle provoquée par des lésions des muscles extra-oculaires. Dans leur expérience de double saut, quand la taille et la direction des sauts initiaux de la cible varient, les auteurs obtiennent une adaptation semblable à celle résultant de la lésion, donc impliquant apparemment les mêmes mécanismes. Selon eux, la lenteur apparente de l'adaptation observée lors de faiblesses musculaires s'expliquerait parce qu'alors toutes les saccades semblent devoir être adaptées indépendamment les unes des autres, quel que soit leur vecteur. En revanche, dans le paradigme de double saut, l'adaptation ne concerne que les saccades effectuées vers les cibles de l'expérience.

Etant donné la diversité des stimulus utilisés, des déplacements des cibles (leurs premiers et/ou seconds sauts) ou des équipements technologiques employés, il n'est pas surprenant que la taille et les décours temporels exacts des adaptations saccadiques rapportées dans la littérature ne soient pas toujours identiques. De plus, à cette disparité entre les protocoles expérimentaux s'ajoute une variabilité inter et même intra-individuelle visible même dans le cadre d'une seule expérience, chez l'être humain (Albano & King, 1989; Frens & van Opstal, 1994; Hopp & Fuchs, 2004, 2010; Pélisson et al., 2010; van Beers, 2007; Wallman & Fuchs, 1998) comme chez le singe (Robinson, Noto, & Bevans, 2003; Straube, Fuchs, Usher, & Robinson, 1997).

Mais le fait que le gain soit modifié par le paradigme de double saut est un effet robuste. Globalement, les expériences qui en un jour font faire à des êtres humains quelques centaines d'essais pendant lesquels la cible effectue un saut intra-saccadique (ou 1000 à 2000 essais à des singes) s'accordent pour conclure que l'adaptation est rarement complète. Cela signifie que ces essais ne modifient pas assez le gain pour que l'amplitude finale des saccades corresponde exactement à la distance séparant le point de fixation et le point d'arrivée de la cible après son second saut. Chez l'être humain, selon Miller et al. (1981) et Semmlow, Gauthier et Vercher (1989), lorsque le saut intra-saccadique de la cible a lieu en direction du point de fixation, le gain peut diminuer de 60 voire 70% de l'amplitude du second saut.

Il reste que la recherche de Scudder et al. (1998) ainsi que la robustesse des effets des sauts intra-saccadiques sur le gain justifient l'emploi du paradigme de McLaughlin (1967) pour étudier les mécanismes adaptatifs du système saccadique, et plus généralement pour investiguer le contrôle de l'amplitude des saccades.

1.2. La nature du signal d'erreur à l'origine de l'adaptation saccadique

Pendant les saccades, non seulement la perception (d'un saut intra-saccadique de la cible par exemple) est entravée, mais encore le contrôle « en temps réel » des déplacements de l'œil est difficile, sinon impossible. Ceci est lié notamment à la vitesse très élevée des saccades (elle peut atteindre 500°/s chez l'être humain selon Becker, 1989, ou Krauzlis, 2008) comparativement à la vitesse de transmission des informations visuelles – 40 à 50 ms, la durée d'une saccade de 10°, est le temps nécessaire pour que l'information visuelle atteigne le colliculus supérieur (Goldberg & Wurtz, 1972; Li & Basso, 2008).

Il est donc communément admis que la précision des saccades repose sur des ajustements qui se font une fois le mouvement terminé et qui se basent sur des informations relatives à

l'inexactitude du comportement effectué (quoique certains auteurs, comme Chen-Harris, Joiner, Ethier, Zee et Shadmehr (2008), Ethier, Zee et Shadmehr (2008) ou encore Gaveau et al. (2003), avancent des arguments en faveur de la modification possible des trajectoires pendant le mouvement même, liée notamment à un changement dans les signaux visuels – comme lorsque la cible effectue un saut intra-saccadique).

Les théories classiques du contrôle moteur ont examiné plusieurs hypothèses sur la nature de ce signal : il pourrait en effet s'agir d'un signal d'erreur dont l'origine est soit motrice, soit visuelle. Afin de tester différentes alternatives, des études ont comparé l'adaptation saccadique obtenue via le paradigme conventionnel de McLaughlin (1967) à celles induites par des procédures visant à manipuler les informations dont disposerait le système.

1.2.1. La réfutation de l'hypothèse d'un signal d'erreur motrice

Deux sortes d'informations relatives à un éventuel signal d'erreur motrice ont été examinées : celles liées à la proprioception et celles provenant des commandes des saccades correctives.

En ce qui concerne les informations proprioceptives venant des muscles extra-oculaires, même si elles sont éliminées par déafférentation, le paradigme de double saut continue à produire des changements de gain (Lewis, Zee, Hayman, & Tamargo, 2001). En revanche, lors de la désadaptation, si les informations proprioceptives sont conservées mais les informations visuelles supprimées (en retardant la réapparition de la cible après son premier saut), le retour au niveau de gain initial est entravé (Seeberger, Noto, & Robinson, 2002). L'implication nécessaire des informations proprioceptives dans le contrôle de l'adaptation saccadique est donc contestée.

Cependant, le contrôle des grandes saccades effectuées alors que les informations visuelles post-saccadiques sont supprimées sembleraient faire exception. En effet Bonnetblanc et Baraduc, (2007) observent une réduction de l'hypométrie, et donc une amélioration de la précision des mouvements, lorsque la cible s'éteint pendant des saccades réalisées vers des cibles situées à 34° d'un point de fixation. Selon les auteurs, le système oculo-moteur se baserait sur ces informations proprioceptives pour compenser une impossibilité de générer des saccades correctives, ces saccades dirigées vers la cible à l'issue des premières saccades lorsque l'image de la cible est située trop loin de la fovéa. Au demeurant, la grande majorité des saccades effectuées quotidiennement sont inférieures à 15° (Bahill, Clark, & Stark, 1975) : cela pourrait expliquer le caractère exceptionnel de cette implication de la proprioception dans le contrôle de si grandes saccades.

Quant au rôle des saccades correctives, il a été démontré qu'elles ne sont pas nécessaires pour que l'adaptation saccadique se produise. Tout d'abord, Wallman et Fuchs (1998) utilisent, chez l'être humain et chez le singe, une modification du paradigme de McLaughlin (1967) : une cible effectue un saut intra-saccadique dans le sens opposé au premier saut, mais elle ne reste que 200 ou 300 ms à cette seconde position avant de retourner à sa première position. Cela réduit considérablement le nombre de saccades correctives dirigées vers l'arrière (recensées dans seulement 1,20% des essais), qui sont par contre observées lors de la majorité des essais du paradigme conventionnel. Dès lors, si le signal d'erreur motrice donné par les saccades correctives était responsable du changement de gain, il ne pourrait pas entraîner de diminution dans cette expérience. Or les résultats montrent bien une réduction du gain saccadique, sans signal d'erreur motrice fourni par les saccades correctives.

Toutefois, cette procédure induit une diminution du gain presque deux fois plus faible que le paradigme conventionnel, testé sur les mêmes singes, dans le même laboratoire. Quelques années plus tard, Noto et Robinson (2001) comparent la diminution du gain induite par le paradigme conventionnel de McLaughlin (1967) et celle obtenue par une procédure légèrement différente, dans laquelle la cible reste visible seulement pendant 90 ms après son second saut avant de disparaître. Cette dernière procédure entraîne aussi peu de saccades correctives que celle de Wallman et Fuchs (1998), par contre elle engendre des changements de gain comparables à ceux provoqués par le paradigme conventionnel.

L'ensemble de ces recherches étayerait donc la position selon laquelle le signal responsable de l'adaptation saccadique ne dépendrait pas des informations relatives à la proprioception ni aux saccades correctives, et qu'il serait plutôt d'origine visuelle .

1.2.2. Un signal d'erreur visuelle à l'origine des changements de gain

a. L'hypothèse d'un signal correspondant à la simple erreur rétinienne post-saccadique

L'erreur visuelle à l'origine des changements de gain correspond à la distance entre l'image de la cible sur la rétine à l'issue de la saccade et la fovéa.

C'est en s'appuyant sur cette définition que l'étude de Noto et Robinson (2001) vise à éclaircir les données de Wallman et Fuchs (1998). Selon Noto et Robinson (2001), lorsque la cible retourne à sa position initiale après son second saut (comme dans l'expérience que Wallman et Fuchs ont réalisée en 1998), une « demande conflictuelle » serait imposée au système oculo-

moteur. En effet, l'erreur post-saccadique est alternativement négative (immédiatement après le saut intra-saccadique de la cible, quand la saccade semble alors hypermétrique) puis positive (lorsque la cible revient à sa position précédente, passant de l'autre côté de la fovéa et rendant la saccade hypométrique). C'est ce qui entraverait la réduction du gain saccadique. En revanche, la procédure de Noto et Robinson (2001) présenterait l'avantage d'éliminer ce conflit, d'où sa capacité à induire une diminution du gain équivalente à ce qui est observé lors du paradigme conventionnel. De plus, les auteurs testent l'effet de cette « demande conflictuelle » : pour la reproduire, ils mêlent des essais d'entraînement où la cible effectue un saut intra-saccadique à des essais où la cible reste sur place. Ils observent alors une diminution du gain significativement moindre.

Ces recherches de Wallman et Fuchs (1998) et Noto et Robinson (2001) les ont amenés à désigner l'erreur rétinienne post-saccadique *per se* comme le signal guidant l'adaptation saccadique. Cela signifie que le système oculo-moteur se baserait sur la lecture directe de cette distance pour modifier le gain de ses saccades en conséquence, réduisant ainsi cette erreur. Cependant il semble que le signal à l'origine de l'adaptation saccadique soit en fait plus complexe.

b. Une définition spatiale de l'erreur parfois ambiguë pour une adaptation saccadique pourtant effective

La définition de l'erreur visuelle en tant que distance rétinienne entre l'image de la cible et la fovéa à la fin de la saccade implique que l'on puisse situer la fovéa par rapport à une stimulation visuelle précise, spécifique et nettement distinguable.

Or cette localisation peut s'avérer ambiguë, comme dans l'étude de Madelain, Harwood, Herman et Wallman (2010). Ces auteurs ont examiné l'influence de deux distances rétiniennes post-saccadiques concomitantes telles qu'elles pourraient imposer un conflit au système oculo-moteur. L'une de ces distances correspond à l'écart entre l'image de la cible et le point d'arrivée de la fovéa à l'issue de la saccade. L'autre est formée par la distance entre l'image d'un distracteur et la fovéa (notons que la demande conflictuelle imposée dans l'expérience de Wallman et Fuchs (1998) vient du fait que la cible apparaît de chaque côté de la fovéa mais de manière successive). Dans une première expérience, la cible effectue un saut intra-saccadique tandis que le distracteur vient la remplacer. Dans une seconde expérience, le distracteur est introduit à la localisation correspondant au saut intra-saccadique tandis que la cible n'effectue pas de second saut. Enfin, Madelain et al. (2010) utilisent le même procédé pour examiner l'influence du distracteur sur la désadaptation (après une adaptation induite via le paradigme conventionnel).

D'une part, les auteurs observent des modifications du gain similaires à celles obtenues par le paradigme conventionnel dans lequel l'erreur rétinienne post-saccadique n'est pas ambiguë. D'autre part, le gain ne change pas si les déplacements de la cible restent constants, même en présence du distracteur. Autrement dit, les changements de gain sont contrôlés uniquement par les déplacements de la cible et non par ceux du distracteur, qui pourtant entraîne lui aussi une stimulation visuelle. Cela conforte le fait que le signal d'erreur responsable de l'adaptation saccadique ne se résume pas à la seule distance rétinienne post-saccadique entre l'image de l'objet et la fovéa.

c. La sélection de la cible, une étape préalable nécessaire

Certaines études indiquent en effet qu'un processus de sélection de la cible constitue un prérequis pour l'adaptation saccadique, ce qui expliquerait notamment les résultats obtenus par Madelain et al. (2010). D'autres données vont dans ce sens.

Ainsi, Noto et Robinson (2005) n'observent pas d'adaptation saccadique chez le singe quand une large image de scène naturelle (de $40^\circ \times 20^\circ$, donc beaucoup plus étendue que le cercle utilisé par Bahcall et Kowler en 2000), dépourvue de cible distincte, effectue un saut intra-saccadique.

Chez l'être humain, Collins, Vergilino-Perez, Beauvillain et Doré-Mazars (2007) examinent l'adaptation de saccades réalisées à l'intérieur d'un même objet (saccades intra-objets) et l'adaptation de saccades émises quand le participant déplace son regard d'un objet à un autre (saccades inter-objets). Dans le premier cas, l'objet, un alignement de 12 X que les participants doivent explorer, mesure 6° de long. Il se déplace pendant la première saccade intra-objet : l'item ayant déjà été atteint, cela n'implique aucun processus de sélection. Dans le second cas, les deux objets correspondent à des alignements de six et cinq X respectivement, totalisant aussi une longueur de 6° . Ils sont tous deux éteints pendant la saccade inter-objets, et seul le second item réapparaît après avoir effectué un saut intra-saccadique. Les participants doivent donc cette fois effectuer une sélection. Les auteurs trouvent que seul le gain des saccades inter-objets est modifié, ce qui suggère aussi que ce processus de sélection de la cible est crucial pour l'adaptation.

d. Un signal responsable de l'adaptation saccadique qui contiendrait des informations relatives au contexte

Il semble que le contexte de l'adaptation saccadique soit un autre élément à prendre en considération dans la définition du signal responsable de l'adaptation saccadique. En effet, en associant des sauts intra-saccadiques positifs à un certain contexte et des sauts négatifs à un autre, il est possible d'induire simultanément une diminution et une augmentation du gain, en fonction de ce contexte. De plus, la différence de gain est visible dès le premier essai consécutif à un changement de contexte.

L'influence de différents types d'indices a ainsi été évaluée. Par exemple, Herman, Harwood et Wallman (2009) montrent que l'adaptation saccadique peut dépendre d'un contexte visuel (une cible illuminée continuellement versus une cible qui clignote). Les changements de gain sont spécifiques aussi de contextes proprioceptifs comme l'orientation de la tête (Shelhamer & Clendaniel, 2002) ou de contextes moteurs tels que la position de l'œil (Alahyane & Pélisson, 2004; Shelhamer & Clendaniel, 2002; Tian & Zee, 2010).

En outre, les données d'Alahyane et Pélisson (2005) suggèrent que l'influence du contexte de l'adaptation saccadique serait durable : après avoir induit une diminution du gain via le paradigme de double saut, ils font passer à plusieurs jours d'intervalle des sessions tests dans lesquelles la cible est éteinte après son premier saut. Cinq jours après la fin de l'expérience, même si le gain commence à remonter, il est encore significativement plus faible que lors de la ligne de base. Pourtant, les participants ont effectué de nombreuses saccades sans double saut entre l'expérience et les différents tests, mais le simple fait d'être replacé dans les mêmes conditions, face au même dispositif expérimental, semble suffire pour que leur gain saccadique reflète l'adaptation saccadique antérieure.

Ces données laissent penser que les participants apprennent une association entre un contexte particulier et un certain niveau de gain et contestent le fait que l'erreur rétinienne post-saccadique en tant que telle puisse être le seul élément qui guide l'adaptation saccadique.

e. Au delà du sens de l'erreur rétinienne : le maintien de l'hypométrie

Certaines expériences révèlent que le sens de cette erreur rétinienne (qui peut-être positive ou négative, selon que la saccade est hypo- ou hypermétrique, respectivement) ne constituerait pas davantage un critère fiable de définition du signal à l'origine de l'adaptation saccadique.

C'est ce que suggère l'une des expériences de Bahcall et Kowler (2000). Leurs participants reçoivent comme consigne de réaliser des saccades (vers la droite et vers la gauche) dont la taille correspond à 75% de la distance séparant le point de départ de la saccade et la cible. Cela signifie que leurs saccades sont toujours hypométriques, que ce soit lors de la phase de pré-adaptation (au début de l'expérience, lorsque la cible ne réalise pas encore de saut intra-saccadique) ou lors de la phase d'adaptation. Bahcall et Kowler (2000) rapportent que le gain change par rapport au niveau de préadaptation quand les sauts intra-saccadiques de la cible sont introduits.

Autrement dit, dans cette situation, le fait que la distance entre l'image de la cible et la fovéa soit toujours positive (du fait de l'instruction) n'entraverait pas la diminution du gain (quand la cible effectue un second saut vers l'arrière) et ne faciliterait pas non plus une augmentation du gain (quand la cible effectue un second saut vers l'avant). Par conséquent, ces résultats semblent aller à l'encontre de l'influence de la seule erreur rétinienne post-saccadique en tant que telle sur l'adaptation saccadique.

Notons cependant que dans cette dernière expérience, qui n'a été passée que par deux participants, et qui à notre connaissance n'a pas été répliquée, les résultats sont entachés d'une variabilité importante. Ainsi, pour le participant naïf, les saccades vers la droite diminuent au fil des essais, même lorsque la cible effectue un saut vers l'avant. Quant aux données de l'autre participant, qui plus est l'un des auteurs, il n'y a pas de différence entre la taille des saccades exécutées lorsque la cible fait un saut vers l'arrière et celle des saccades effectuées lorsqu'il n'y a pas de saut intra-saccadique. En outre, il est difficile de savoir la part des résultats qui relève des instructions.

L'examen de l'hypométrie « naturelle » des saccades fournit d'autres éléments montrant que le signal contrôlant le gain saccadique ne peut se résumer à la seule erreur rétinienne.

En effet, l'amplitude des saccades observées en laboratoire correspond à 90% de l'excentricité de la cible en moyenne (Becker, 1989; Irving, Steinbach, Lillakas, Babu, & Hutchings, 2006). A ce jour, la raison exacte n'en est pas encore connue. Certains auteurs, comme Robinson et al. (2003) et Havermann et Lappe (2010) postulent l'existence d'un mécanisme qui maintiendrait activement l'hypométrie initiale des saccades. Ils explorent, respectivement chez le singe et l'être humain, les changements de gain des saccades induits par une cible qui effectue des sauts intra-saccadiques définis par rapport à la position de l'œil à la fin de la saccade. En l'occurrence, la cible réapparaît à moins de 0.5° à côté de la position d'arrivée de l'œil, mesurée en temps réel. Si l'erreur rétinienne post-saccadique était la seule information prise en compte par le système, aucun changement de gain significatif ne devrait se produire avec

de tels déplacements de cible. Or Robinson et al. (2003) et Havermann et Lappe (2010) observent une réduction du gain dépassant le seuil de significativité.

Ce postulat d'un tel mécanisme repose notamment sur une expérience que Henson a réalisée en 1978, dans laquelle il utilise un dispositif optique de telle sorte qu'à la fin des saccades, la cible apparaisse 2 ou 3° plus près du point de fixation initial. Henson (1978) observe que le pourcentage de saccades hypométriques revient à son niveau initial en 10 minutes environ.

Toutefois, même si un tel mécanisme existait, il ne suffirait pas pour rendre compte de l'adaptation saccadique. Par exemple, dans le paradigme princeps de McLaughlin (1967) où la cible effectue un saut intra-saccadique vers l'arrière, il semble paradoxal que la diminution du gain ne soit que rarement équivalente à la taille du saut intra-saccadique (Hopp & Fuchs, 2004). Dans le même ordre d'idée, lorsqu'une petite erreur rétinienne constante est imposée au système, le fait que l'adaptation saccadique atteigne une asymptote reste inexpliqué (Robinson et al., 2003) : si le mécanisme maintenant l'hypométrie était le seul en jeu, l'adaptation ne devrait pas s'arrêter. En outre, une autre contradiction apparaît dans les résultats de Havermann et Lappe (2010), car chez trois des sept participants de cette condition expérimentale où les sauts intra-saccadiques sont minimes, le gain saccadique augmente.

En bref, même si l'existence de ce mécanisme dont le rôle serait de maintenir activement l'hypométrie initiale reste hypothétique, ces trois études témoignent elles aussi que la simple lecture et la réduction de l'erreur rétinienne post-saccadique en tant que telle ne sont pas les seuls éléments impliqués dans le contrôle du gain saccadique.

1.2.3. Un signal élaboré à partir de la comparaison entre l'erreur prédite et l'erreur réelle

Pour tenter de définir le signal qui contrôle l'adaptation saccadique, Bahcall et Kowler (2000) invoquent un « modèle de comparaison visuelle ». Selon eux, le système oculo-moteur comparerait l'image rétinienne réelle, obtenue à l'issue de la saccade, et l'image qui était prédite à partir de la taille et de la direction de la saccade imminente. Une importante disparité entre ces deux images générerait le signal d'erreur provoquant l'adaptation saccadique, le changement du gain se poursuivant jusqu'à ce que l'image visuelle post-saccadique corresponde à la prédiction.

L'un des arguments de Bahcall et Kowler (2000) s'appuie sur l'effet délétère qu'un délai de 400 ms, introduit entre la fin de la saccade et la réapparition de la cible à sa seconde position, exerce sur l'adaptation saccadique. Ces constatations sont corroborées deux ans plus tard par les

données de Fujita, Amagai, Minakawa et Aoki (2002) qui observent une corrélation inverse entre le changement du gain saccadique et le délai de réapparition de la cible. Puisque ces paramètres temporels n'altèrent pas les caractéristiques spatiales, rétinienne, de l'erreur visuelle, ils sembleraient plutôt affecter les processus plus élaborés qui seraient, eux, à l'origine du signal d'erreur. Ces processus sous-jacents nécessiteraient donc que les informations visuelles soient disponibles rapidement à la fin du mouvement.

Par ailleurs, selon Bahcall et Kowler (2000), ce modèle de comparaison expliquerait aussi leurs autres résultats, tels l'adaptation saccadique induite par des cibles plus larges ou plus complexes que les cibles isolées généralement utilisées. Il permettrait de rendre compte également des changements de gains observés quand les participants ont pour tâche de réaliser des saccades hypométriques, et donc que l'erreur rétinienne est toujours positive – la prédiction jouant alors un rôle prépondérant.

Dans une expérience récente, Wong et Shelhamer (2011) tirent profit de l'hypométrie des saccades pour dissocier l'erreur visuelle et l'erreur prédite sans donner comme consigne de faire des saccades trop courtes. Ils imposent au système une petite erreur post-saccadique constante : la cible effectue des sauts intra-saccadiques l'amenant à 0.7° de la fovéa, toujours du côté opposé au point de fixation (elle arrive au point C2 de la figure 3). L'erreur rétinienne (la distance entre le point d'arrivée de la cible et le point d'arrivée de la fovéa) est donc constamment positive. Si celle-ci était la seule à guider l'adaptation saccadique, le gain devrait augmenter, or les chercheurs observent une réduction du gain saccadique dans 82% des cas.

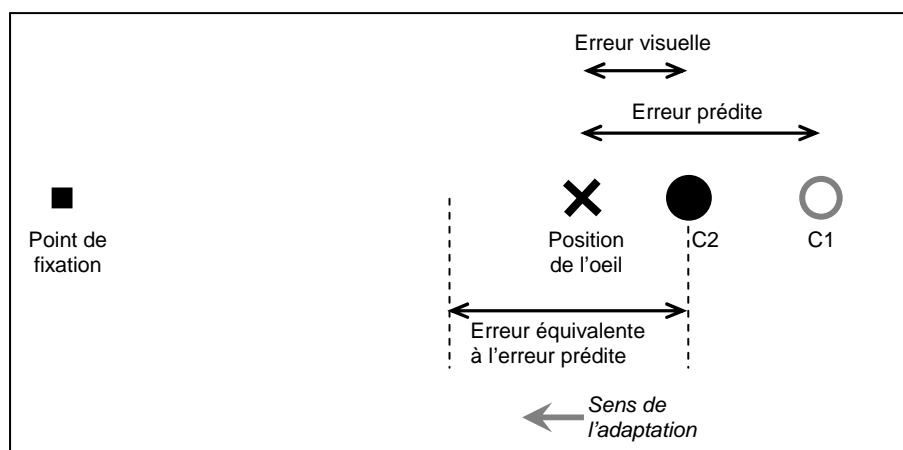


Figure 3. Schéma de la procédure de Wong et Shelhamer (2011). C1 : position de la cible après son premier saut, C2 : position de la cible après son saut intra-saccadique constant de 0.7° .

Wong et Shelhamer (2011) interprètent leurs résultats en termes d'erreur de prédiction. Plus précisément, ils postulent que l'hypométrie initiale (de 1.63° en moyenne chez leurs participants) correspondrait à l'erreur prédite par le système. Par conséquent, le signal guidant l'adaptation saccadique serait défini par la différence entre cette hypométrie et l'erreur observée (0.7°) imposée dans cette expérience. Cette différence (de 0.93° en moyenne) expliquerait la réduction du gain en direction du niveau de l'hypométrie initiale.

De façon intéressante, Wong et Shelhamer (2011) observent que moins l'hypométrie initiale est marquée, plus le gain est réduit par la procédure. Non seulement cela appuie leur hypothèse, mais cela permet aussi de rendre compte de l'augmentation du gain obtenue chez deux de leurs participants ainsi que dans une expérience contrôle où les sauts intra-saccadiques sont de 2° . En effet, dans ces derniers cas l'hypométrie initiale est plus faible que l'erreur observée consécutivement aux sauts de la cible. Cela changerait donc le sens de la différence entre l'erreur prédite et l'erreur observée.

Selon Wong et Shelhamer (2011), ces résultats étayeraient donc l'existence de modèles internes responsables de ces comparaisons, et donc de l'élaboration du signal responsable de l'adaptation saccadique.

1.3. Deux modèles internes *adaptatifs* pour rendre compte de l'adaptation saccadique

1.3.1. L'implication simultanée de mécanismes distincts

Un faisceau de données empiriques récentes semble soutenir l'hypothèse selon laquelle différents mécanismes seraient impliqués dans l'adaptation saccadique (Semmlow et al., 1989). Notamment, dans leur étude computationnelle, Ethier et al. (2008) simulent et comparent les trajectoires des saccades après adaptation saccadique (induisant une augmentation ou une diminution du gain) aux trajectoires de saccades contrôles, c'est-à-dire des saccades de même amplitude mais émises vers des cibles stationnaires. Dans la mesure où ils observent une modification des propriétés des saccades dont l'amplitude a été diminuée par adaptation saccadique, ils proposent que la réduction du gain serait spécifiquement liée à l'adaptation de modèles internes. En revanche, quand l'adaptation saccadique vise à augmenter le gain, les propriétés des saccades restent similaires à celles des saccades contrôles. Ethier et al. (2008) en

infèrent que l'augmentation du gain serait quant à elle le corollaire d'une transformation de la représentation des coordonnées de la cible (un processus nommé « *remapping* » en anglais).

Des études réalisées chez le singe suggèrent aussi que des mécanismes distincts responsables de l'adaptation saccadique évolueraient de manière parallèle. Ainsi, Kojima, Iwamoto et Yoshida, (2004) induisent une adaptation saccadique dans un sens, puis ils inversent la direction des sauts intra-saccadiques pour produire une désadaptation, qu'ils font immédiatement suivre d'une nouvelle adaptation. Ils observent une accélération des changements lors de la réplication de l'adaptation. Cette facilitation s'observe aussi si les singes sont plongés dans l'obscurité entre la désadaptation et la seconde phase d'adaptation. Ensuite, lors de ces réplifications, le taux de changement du gain finit par ralentir. Les auteurs interprètent ces phénomènes de facilitation et de ralentissement comme les modifications concomitantes de deux mécanismes sous-jacents, l'un responsable de la diminution du gain, l'autre de l'augmentation.

Les mécanismes impliqués dans l'adaptation saccadique ne semblent pas nécessairement fonctionner selon des échelles de temps identiques. Robinson, Soetedjo et Noto (2006) appliquent le paradigme de double saut chez trois singes, auxquels ils font faire une session d'adaptation saccadique quotidienne pendant 19 jours (en vue de faire diminuer le gain saccadique). Après chaque session, les animaux sont placés dans l'obscurité. A la fin de la première session d'adaptation, le pourcentage de diminution du gain correspond en moyenne à une adaptation de 61% de ce qui est attendu. En revanche, à la fin du 19^{ème} jour, l'adaptation est quasiment complète. En outre, Robinson et al. (2006) notent une réduction progressive puis la disparition de la désadaptation transitoire observée entre deux journées. Robinson et al. (2006) en infèrent la co-influence de deux mécanismes, l'un évoluant à court terme et l'autre à long terme.

Körding et al. (2007) ont formalisé cette hypothèse à l'aide de la théorie bayésienne. Ils postulent que le système oculo-moteur, lorsqu'il est confronté à des erreurs de performance, s'appuierait sur les représentations qu'il se fait des perturbations (ainsi que de la variabilité de ses propres conclusions) pour modifier le gain saccadique en conséquence. Le gain reflèterait le poids attribué à ces perturbations en termes d'échelle de temps. Selon ce raisonnement, plus les perturbations paraîtraient transitoires au système (telles que celles qu'il observe au commencement des expériences de double saut par exemple), plus vite elles seraient oubliées, et plus fugace serait leur effet sur le gain. En revanche, à la fin des expériences, ou après les 19 jours de l'expérience subie par les singes de Robinson et al. (2006), le système attribuerait ces perturbations à des phénomènes plus durables, les effets seraient alors conservés plus longtemps en mémoire, cela se répercutant aussi sur le gain.

Cette attribution expliquerait non seulement les effets de facilitation observés par Kojima et al. (2004), mais aussi l'adaptation complète ainsi que la disparition de la désadaptation transitoire notées par Robinson et al. (2006). En outre, le système devrait également décider de l'origine des perturbations. En effet, les changements peuvent provenir soit de l'environnement, soit du système lui-même. Selon Wong et Shelhamer (2011), les changements que le système attribuerait à des causes internes l'inciteraient à altérer ses commandes motrices. En revanche, des perturbations que le système considérerait comme provenant de l'environnement l'amèneraient davantage à modifier ses prédictions (c'est-à-dire à s'attendre à une erreur). Ces attributions ne se faisant pas en tout ou rien, cela expliquerait notamment le fait que, dans leur expérience de 2011 comme dans celle de Robinson et al. (2003), le gain atteigne une asymptote alors que l'erreur rétinienne post-saccadique reste constante.

En résumé, ces études suggèrent que l'adaptation saccadique relèverait de l'action conjuguée de mécanismes internes qui non seulement seraient distincts, mais qui évolueraient aussi selon des échelles de temps qui leur seraient propres.

1.3.2. Deux modèles internes pour rendre compte du contrôle moteur

Les théories du contrôle optimal ont formalisé cette approche. Selon elles, deux modèles internes expliqueraient le contrôle et l'apprentissage sensori-moteurs (ex. : Wolpert & Ghahramani, 2000; Wolpert, Ghahramani, & Flanagan, 2001; Wolpert, Ghahramani, & Jordan, 1995).

Les modèles internes correspondraient aux représentations qu'élaborerait le système nerveux central relativement aux transformations des informations sensorielles en commandes motrices et réciproquement. Ces modèles renseigneraient sur les différents états du système : état désiré (but), état prédit, état actuel. Un état correspond à l'ensemble des informations relatives aux propriétés du corps (en l'occurrence l'activation des muscles, la position et la vitesse des yeux) et de l'environnement (comme la position de la cible), des données qui sont nécessaires pour pouvoir prédire et contrôler le mouvement. Les principales interactions entre ces modèles telles qu'elles sont conçues par les théories computationnelles sont évoquées ci-après, de manière succincte et non exhaustive, et sont schématisées par la figure 4.

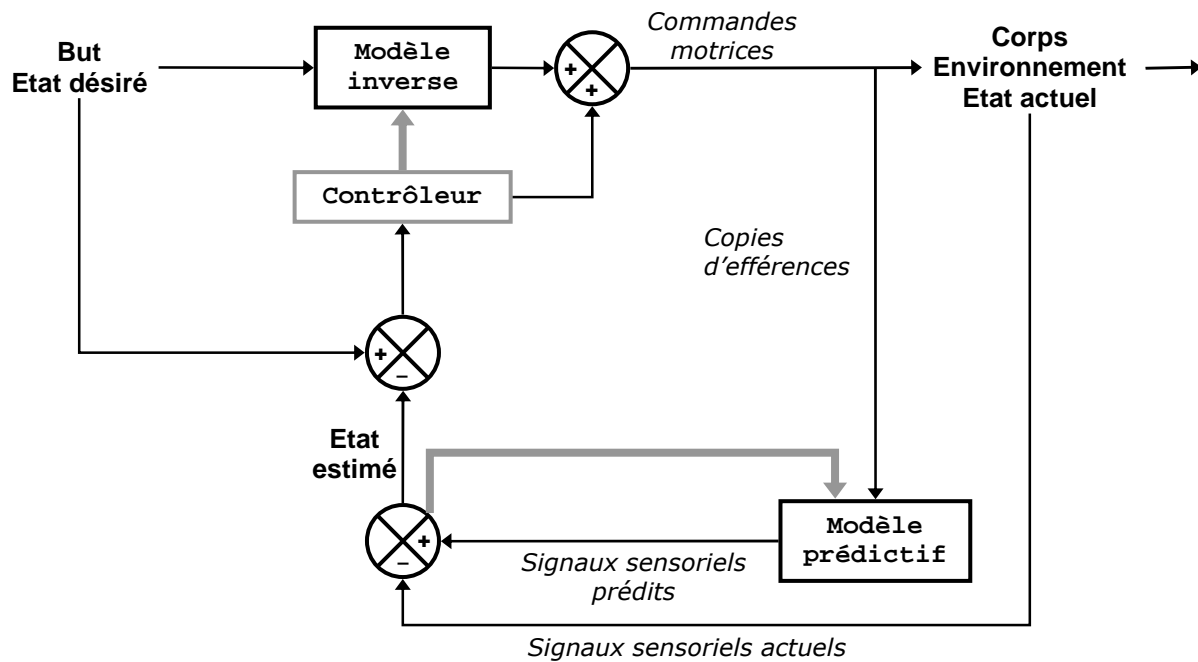


Figure 4. schéma simplifié du contrôle sensori-moteur tel qu'il est envisagé par les théories computationnelles. En gris sont représentés les éléments spécifiquement impliqués dans l'adaptation des modèles internes.

L'un de ces modèles, appelé modèle inverse (« *inverse model* »), permettrait au système de se représenter la transformation consistant à passer des informations sensorielles désirées (correspondant au but de l'action, par exemple, distinguer une lettre particulière dans un mot) aux commandes motrices nécessaires pour réaliser ces actions (une saccade oculaire) – et d'atteindre ce but.

Ce modèle inverse transformerait continuellement l'état désiré en commandes motrices, pendant toute la durée du mouvement. Pendant ses actions, l'état du système change, provoquant des conséquences sensorielles précises. A priori, le système aurait besoin de ces signaux sensoriels pour éventuellement corriger ses mouvements. Mais en général, obtenir ces informations sur l'état actuel pose deux problèmes : d'une part, les informations sensorielles mettent plus de temps pour parvenir au système qu'il n'en faut pour exécuter une saccade (Goldberg & Wurtz, 1972; Li & Basso, 2008), d'autre part elles peuvent être incertaines à cause de la présence de bruit neuronal (Faisal, Selen, & Wolpert, 2008).

Afin de pallier ces problèmes, le système sensori-moteur s'appuierait sur des prédictions. L'autre modèle interne, le modèle prédictif (« *forward model* »), entrerait ici en jeu. Il se baserait sur les copies des commandes motrices, ou copies d'efférences. Selon les théories computationnelles actuelles, il permettrait au système de se représenter les relations causales entre les commandes motrices et les conséquences sensorielles des actions – la transformation inverse du modèle précédent. Autrement dit, grâce à ce modèle interne, le système sensori-moteur pourrait prédire les conséquences sensorielles actuelles avant que les informations sensorielles lui parviennent en retour, avant même que le mouvement soit terminé. Le système comparerait ces estimations avec les informations sensorielles désirées, et il serait alors en mesure de corriger éventuellement ses commandes motrices ultérieures.

Certaines données expérimentales peuvent en effet amener à inférer l'existence des copies d'efférences, et leur implication dans le contrôle des saccades oculaires. Par exemple, Guthrie, Porter et Sparks (1983) étudient le contrôle des saccades en réponse à des microstimulations chez des singes dont les muscles extra-oculaires ont été déafférentés. Les singes doivent émettre des saccades vers des cibles affichées pendant 50 ms. Lors de certaines saccades, les auteurs stimulent le colliculus supérieur des animaux juste après la disparition de la cible, ce qui entraîne un mouvement des yeux vers une nouvelle localisation. Les singes émettent alors des mouvements compensatoires leur permettant d'atteindre la cible de manière aussi précise que lorsqu'il n'y a pas de perturbation. Les auteurs déduisent qu'en l'absence d'informations proprioceptives et visuelles (la cible ayant disparu), ce serait la copie d'efférence qui permettrait cette compensation.

1.3.3. Deux modèles internes adaptatifs

Les théories computationnelles formalisent l'adaptation saccadique comme étant liée à l'adaptabilité de chacun des deux modèles internes. En d'autres termes, elles supposent que le modèle inverse comme le modèle prédictif seraient en mesure de modifier leurs propres informations de sortie (Wolpert & Ghahramani, 2000; Wolpert et al., 2001). Il semble que ces adaptations puissent avoir lieu en quelques minutes – le temps mis par la procédure de double saut pour altérer le gain des saccades.

D'une part, les théories computationnelles postulent que le modèle inverse changerait les relations qu'il établit entre les conséquences sensorielles désirées et les commandes motrices générées (ses sorties).

Elles doivent alors conceptualiser la manière dont ce modèle pourrait avoir accès au signal d'erreur. En effet, le signal de ce modèle interne devrait être de nature motrice (lorsque les commandes motrices ne permettent pas d'atteindre l'état désiré), alors que les informations dont ce modèle dispose sont de nature sensorielle. L'existence d'un contrôleur (« *feedback controller* ») a été alors invoquée (Wolpert & Ghahramani, 2000; Wolpert et al., 2001). Celui-ci serait capable d'opérer la traduction du signal d'erreur (la différence entre l'état estimé et l'état désiré) en coordonnées motrices. Les informations issues de ce contrôleur viendraient, le cas échéant, s'ajouter aux commandes motrices générées par le modèle inverse. L'absence de telles informations signifierait une adéquation parfaite entre ces commandes motrices et l'état désiré. De plus, les informations issues du contrôleur parviendraient également au modèle inverse afin de lui permettre de corriger ses sorties ultérieures.

D'autre part, lors de ces expériences de double saut, le modèle prédictif s'adapterait en se basant sur un signal d'erreur de nature sensorielle, correspondant à la différence entre les conséquences sensorielles prédites et les signaux sensoriels réels.

Dans le cas de mouvements relativement lents, comme l'atteinte ou la saisie d'un objet, ces signaux proviennent des différentes modalités sensorielles auxquelles le système a accès à l'issue du mouvement, mais aussi de celles dont il dispose pendant le mouvement. Grâce à cela, le système serait capable d'ajuster les trajectoires en temps réel. En effet, les deux modèles internes seraient à l'œuvre pendant toute la durée du déplacement. D'un côté, le modèle inverse générerait des commandes motrices correspondant à l'état désiré en tout point de la trajectoire. De l'autre côté le système intégrerait continuellement les prédictions internes et les signaux sensoriels réels afin d'obtenir les estimations les plus précises possibles des conséquences de ses actions (Wolpert & Ghahramani, 2000; Wolpert et al., 2001; Wolpert et al., 1995).

En revanche, dans le cas des mouvements extrêmement rapides que sont les saccades oculaires, les informations ne semblent pas pouvoir parvenir au système pendant le mouvement. Les saccades sont souvent décrites comme balistiques, dans le sens où leur trajectoire est programmée avant leur réalisation : selon la plupart des théoriciens, le contrôle visuel en temps réel serait impossible. Les ajustements des commandes motrices en vue des saccades ultérieures auraient lieu uniquement à partir des signaux sensoriels obtenus à l'issue du mouvement, c'est-à-dire à partir de l'erreur de position rétinienne post-saccadique – que la procédure de double saut engendre artificiellement (ex. : Bonnetblanc & Baraduc, 2007; Collins & Wallman, 2011; Harris, 1995; Harris & Wolpert, 1998, 2006; Lewis et al., 2001; van Beers, 2008; Wong & Shelhamer, 2011).

Notons que les chercheurs de l'équipe de Shadmehr postulent, eux, que même s'il ne semble pas exister de contrôle visuel en temps réel pendant les saccades, le système oculo-moteur serait en mesure de modifier la trajectoire au cours du déplacement (Chen-Harris et al., 2008; Ethier et al., 2008). Pour cela, le système s'appuierait sur les estimations du modèle prédictif, estimations qui seraient disponibles pendant toute la durée du mouvement. Dès lors, ils font l'hypothèse que l'éventuelle capacité des modèles internes à s'adapter aux changements de l'environnement devrait se refléter au niveau des trajectoires des saccades, trajectoires qu'ils simulent en fonction de l'implication (ou non) de chacun des deux modèles internes.

Dans cette perspective, Chen-Harris et al. (2008) modifient la procédure conventionnelle de double saut de McLaughlin (1967) : après un saut initial horizontal, leurs cibles effectuent un saut intra-saccadique vertical, dès que l'œil s'éloigne du point de fixation. Les auteurs analysent l'évolution des trajectoires des saccades en les divisant en quatre segments dont ils mesurent la courbure. Ils observent que la courbure augmente au fil des essais de manière plus marquée pour le dernier segment, surtout par rapport au premier segment. Ces trajectoires sont cohérentes avec la simulation d'une adaptation plus rapide du modèle prédictif que du modèle inverse.

Chen-Harris et al. (2008) apportent ainsi un argument computationnel en faveur de leurs hypothèses : d'une part le système oculo-moteur, se basant sur le modèle prédictif, serait en mesure de corriger les trajectoires des saccades en temps réel, et d'autre part les deux modèles internes s'adaptent selon des échelles de temps différentes. Selon eux, l'influence du modèle prédictif semblerait prépondérante dans ce type d'adaptation saccadique : une recalibration des effets des copies d'efférences viserait notamment à changer les estimations de ce modèle.

Il reste que leur conclusion se fonde sur deux suppositions. D'une part, rien ne prouve que les premiers et derniers segments des trajectoires reflètent effectivement les sorties respectives des modèles internes et prédictifs. D'autre part, Chen-Harris et al. (2008) considèrent la courbure des trajectoires comme la signature des changements se produisant au sein des modèles, or certaines données pourraient le remettre en question. En effet, certaines recherches indiquent qu'un traitement des informations visuelles pendant la saccade elle-même n'est pas impossible.

Ainsi, Gaveau et al. (2003) montrent que lorsqu'un saut intra-saccadique a lieu une fois toutes les dix saccades, donc en dehors de tout phénomène d'adaptation, l'amplitude et le pic de vitesse de ces saccades diffèrent significativement (de 1.2° et $14.2^\circ/\text{s}$ en moyenne) de ceux des saccades contrôles exécutées lorsque la cible est stationnaire. Ces deux mesures augmentent lorsque les sauts intra-saccadiques sont positifs et elles diminuent lorsque ces sauts sont négatifs. Puisque les sauts intra-saccadiques sont peu fréquents, ces changements ne semblent pas relever de la modification des prédictions de modèles internes, mais plutôt à l'influence d'un signal rétinien.

En outre, les données de Panouillères, Urquizar, Salemme et Pélisson (2011) sur la durée minimale de la présence de la cible nécessaire pour induire l'adaptation saccadique dans le paradigme conventionnel de double saut pourraient également aller dans ce sens. Les auteurs repèrent le début de la saccade pour afficher la cible à sa seconde localisation ; celle-ci est donc présente 55 ms environ avant de disparaître, dont 40 ms correspondent à la durée du mouvement. Dans cette situation, l'adaptation saccadique n'est pas différente de celle obtenue lorsque la cible est affichée plus longtemps. Soit les 15 ms pendant lesquelles la cible est visible après l'achèvement de la saccade principale suffiraient au système pour élaborer son signal d'erreur, soit le système serait effectivement capable de traiter le saut intra-saccadique pendant le mouvement.

L'étude de Gaveau et al. (2003) indique aussi que des sauts intra-saccadiques occasionnels, s'ils modifient l'amplitude et la vitesse des saccades, ne permettent pas d'induire des changements comparables à ceux obtenus lors de la répétition de ces doubles sauts.

Mais, que le système s'appuie sur des signaux visuels obtenus pendant ou à l'issue du mouvement, les changements des modèles internes auraient lieu sur des échelles de temps assez courtes (de l'ordre de la dizaine de minutes) pour pouvoir se manifester au cours du paradigme expérimental de double saut. C'est par ces processus adaptatifs que les théories computationnelles rendent compte en partie de l'apprentissage au sein du système saccadique.

1.4. Le système oculo-moteur, un système adaptatif qui « apprendrait à effectuer des mouvements optimaux »

Selon les théories du contrôle optimal, si atteindre les cibles visuelles correspond généralement au but des saccades oculaires (autrement dit si la projection de l'image de la cible sur la fovéa correspond aux conséquences sensorielles désirées), le système tenterait en outre de le faire de manière optimale. Cet autre aspect de l'apprentissage oculo-moteur, qui relèverait de l'ontogenèse, serait visible sur une échelle de temps plus longue que celle de l'adaptation.

Pour étudier ce type d'apprentissage au sein du système oculo-moteur, les auteurs ne se sont pas appuyés sur des paradigmes tels que l'adaptation saccadique, mais sur plusieurs constats, comme l'existence de régularités dans les paramètres des mouvements.

1.4.1. Les constats initiaux : des régularités en dépit du bruit interne

Il est admis depuis longtemps que les caractéristiques dynamiques des saccades sont stables. Les relations stéréotypées entre la durée de la saccade et l'amplitude d'une part, et entre le pic de vitesse et l'amplitude d'autre part, ont très souvent été constatées, elles sont connues sous le nom de « *main sequence* » (ex. : Bahill et al., 1975; van Beers, 2007) – nous verrons ultérieurement dans quelle mesure ce caractère stéréotypé peut toutefois être remis en question. Une seconde régularité concerne le gain des saccades : les études relevant une hypométrie en laboratoire abondent (ex : Becker, 1989; Harris, 1995; Havermann & Lappe, 2010; Henson, 1978; Irving et al., 2006; Robinson et al., 2003; Wong & Shelhamer, 2011).

Selon les théories du contrôle optimal, la *main sequence* ainsi que cette tendance générale à l'hypométrie seraient le reflet de la capacité du système à réaliser des calculs l'amenant à effectuer certains mouvements – les mouvements optimaux – plutôt que d'autres. Ces calculs portent sur les différents coûts associés aux mouvements. En effet, chaque déplacement est caractérisé par une dépense énergétique, une stabilité, une durée et une précision spécifiques ; par exemple, un mouvement lent est souvent plus précis qu'un mouvement rapide. Des compromis entre les différents coûts sembleraient être réalisés ; autrement dit le système nerveux central générerait des mouvements de manière à ce que des fonctions de coûts spécifiques soient minimisées (Wolpert & Ghahramani, 2000; Wolpert et al., 2001)

Dans le cas des saccades oculaires, les muscles extra-oculaires étant résistants à la fatigue (Fuchs & Binder, 1983) et les mouvements très simples, contrôlés par trois paires de muscles, il est difficile d'imaginer que les coûts métaboliques et ceux liés à une éventuelle instabilité soient élevés ; ces derniers ne rentrent donc pas dans l'équation. Selon les théories du contrôle optimal, le système saccadique semblerait donc devoir essentiellement gérer un compromis entre vitesse et précision. Cette idée a été décrite notamment dans les travaux de Harris, Wolpert et leurs collaborateurs.

Ces auteurs partent du constat qu'il existe un bruit interne, aléatoire, incontrôlable, à différents niveaux du système nerveux (voir Faisal et al., 2008, pour une revue, dont est tirée la figure 5 qui donne un aperçu des sources possibles de bruit neuronal). Ce bruit affecte chaque étape des comportements sensori-moteurs : de la sensation jusqu'à la réponse motrice, en passant par la perception, le traitement de l'information, la planification et l'exécution du mouvement. C'est de ce bruit interne que résulte une partie de la variabilité observée dans les comportements, variabilité que l'on constate aussi dans les mouvements oculaires (van Beers, 2007).

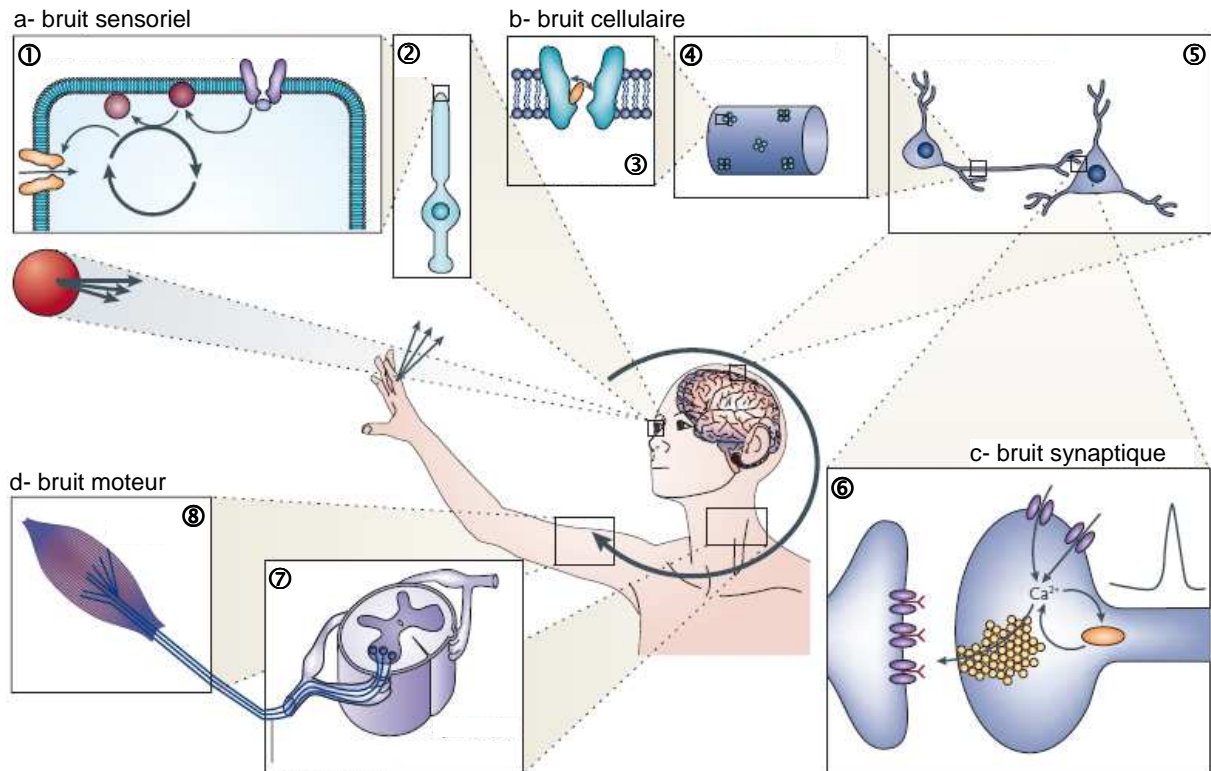


Figure 5. Aperçu de plusieurs sources possibles de bruit interne lors des différentes étapes d'un comportement sensori-moteur. (source : Faisal et al., 2008.)

1.4.2. La *main sequence*, ou la signature d'un mouvement optimal

En 1995, Harris propose un modèle pour rendre compte de l'hypométrie quasi systématique des saccades. Elle serait liée à deux paramètres. Le premier, la durée de la saccade, conférerait un coût au mouvement dans la mesure où la vision est dégradée pendant le déplacement de l'œil. Par conséquent, le système éviterait les saccades hypermétriques car, en augmentant la distance parcourue par l'œil et donc la durée du déplacement (les deux propriétés étant corrélées), elles augmentent le temps pendant lequel la vision est extrêmement réduite. La seconde contrainte serait liée au manque de précision des saccades : le gain moyen ne doit pas être trop éloigné de 1 sous peine de compromettre la vision nette de la cible à l'issue de la saccade, ou d'entraîner des saccades correctives, ce qui s'avère coûteux également.

Par ailleurs, le gain moyen des saccades est assorti de la variabilité évoquée ci-dessus (van Beers, 2007), inhérente à tout comportement. Le système devant compenser cette variabilité, il

réduirait le gain saccadique moyen afin de minimiser la proportion de saccades hypermétriques, mais jusqu'à une certaine limite imposée par la nécessité de la précision. Les simulations de saccades que Harris (1995) obtient en tenant compte de ces paramètres sont cohérentes avec la *main sequence* caractéristique des saccades effectuées par l'adulte humain : cela appuie l'idée d'un gain optimal (inférieur à 1) tel que le temps passé à réaliser une seule saccade soit minimisé étant donné la variabilité des erreurs de position.

Quelques années plus tard, Harris et Wolpert (1998, 2006) revisitent ce modèle. Ils partent de l'axiome que le bruit neuronal est proportionnel à la taille du signal à l'origine des saccades. En présence d'un tel bruit, la répétition des mouvements effectués pour une amplitude donnée va donc produire une certaine distribution d'erreurs post-saccadiques – une certaine variabilité. Plus le mouvement programmé est rapide, plus l'activation neuronale augmente, et plus cette variabilité est élevée – plus les saccades sont imprécises.

La solution que le système pourrait adopter pourrait consister à ralentir ses mouvements. Or une seconde contrainte pèse sur le système : celle qui est liée à la durée du mouvement. Cette fois, plus le mouvement est lent, plus le temps passé avec une vision dégradée est long. Le problème auquel le modèle de Harris et Wolpert (1998, 2006) répond consiste à trouver le mouvement optimal, c'est-à-dire la durée optimale pour une amplitude saccadique donnée, qui permette au système de résoudre la tension entre ces deux coûts (le « *speed-accuracy trade-off* »).

En définissant le coût total optimal d'une saccade comme le minimum de ces deux fonctions de coûts, et en simulant ces coûts pour des saccades de différentes amplitudes, le modèle de Harris et Wolpert (1998, 2006) prédit que ces durées optimales vont augmenter avec l'amplitude. La fonction d'optimisation obtenue prédit effectivement la *main sequence*, cette relation stéréotypée entre amplitude, durée et pic de vitesse largement attestée de manière empirique (ex : Bahill et al., 1975). Ce modèle constitue le point de départ de bon nombre d'études computationnelles qui abondent dans ce sens (ex. : van Beers, 2008; Wolpert et al., 2001).

En résumé, les théories computationnelles considèrent que le signal responsable de l'adaptation saccadique consisterait en un signal complexe, extrêmement élaboré, dérivé de l'erreur rétinienne post-saccadique.

Face aux perturbations infligées au système oculo-moteur lors des expériences de double saut, seraient impliqués des modèles internes spécialisés, en constante interaction. Grâce à ces mécanismes adaptatifs responsables notamment de la recalibration de gains internes, le système

conserverait sa capacité à établir la correspondance entre l'excentricité de la cible et la force musculaire nécessaire à l'œil pour l'atteindre du regard.

En outre, les saccades ont également permis à ces théoriciens d'investiguer l'apprentissage sensori-moteur à plus long terme, au niveau ontogénétique. En particulier, les théories du contrôle optimal défendent l'idée que le système oculo-moteur apprendrait à optimiser des fonctions de coût. La *main sequence* apparaîtrait selon elles comme une stratégie que le système adopterait pour optimiser la vision en présence d'une variabilité délétère.

1.5. Quelques limites des théories computationnelles du contrôle moteur

Les théories computationnelles du contrôle moteur proposent des outils performants, c'est-à-dire capables de bien rendre bien compte des données expérimentales (Faisal et al., 2008; Glimcher, 2003; Karniel, 2011; Körding, 2007; Wolpert & Ghahramani, 2000; Wolpert et al., 2001). Cependant leurs modèles se heurtent à certaines limites et les théoriciens ne s'accordent pas toujours entre eux.

Ainsi, en 2011, Karniel a écrit une revue de littérature sur quelques unes de leurs controverses : il relève par exemple que certains théoriciens postulent l'existence de deux modèles internes, d'autres un seul, d'autres encore (comme Wolpert & Kawato, 1998) avancent que plusieurs modèles pourraient être impliqués dans le contrôle moteur. Karniel (2011) évoque également les questions de la redondance ou de la représentation du temps (du passé, du futur, comme du temps dans lequel s'inscrit les actions), qui ne sont pas tranchées.

En outre, les recherches sur le contrôle des saccades présentées précédemment montrent que les théoriciens ne s'accordent pas non plus sur la manière dont le contrôle a lieu : les saccades sont-elles balistiques (Harris & Wolpert, 1998, 2006; van Beers, 2008), ou les trajectoires peuvent-elles être contrôlées pendant le mouvement (Chen-Harris et al., 2008; Ethier et al., 2008) ? Par ailleurs, quels coûts impliquer, quel critère d'optimisation utiliser dans les modélisations ? Si certains modèles impliquent les coûts liés à la nécessité de la précision et à la vitesse des mouvements (Harris & Wolpert, 1998, 2006; van Beers, 2008), d'autres intègrent celui de l'effort (Shadmehr, Orban de Xivry, Xu-Wilson, & Shih, 2010). Ce bref aperçu de quelques contradictions ne prétend pas être exhaustif.

Au-delà de ces discordances qui portent sur la manière dont ces théories conceptualisent le contrôle moteur, d'autres questions d'ordre épistémologique peuvent être soulevées.

Avec son modèle inverse en guise d'unité centrale, sa boucle de rétroaction servant à réguler les sorties (les commandes motrices) dans le but de réaliser la concordance entre une intention et ses réalisations, le système oculo-moteur tel qu'il est conceptualisé par la majorité des théories computationnelles peut être comparé à un servomécanisme (Ingvaldsen & Whiting, 1997).

Dans la mesure où ces modèles parviennent à bien décrire les comportements, cette métaphore du servomécanisme semblerait de prime abord particulièrement bien choisie. Mais c'est justement parce qu'elles utilisent une métaphore que ces théories peuvent susciter un certain scepticisme.

Qu'il s'agisse d'un dispositif électronique (tel que le thermostat d'un chauffage central) ou d'un – putatif – système biologique, le problème fondamental est de trouver les règles de transformations. En effet, ces deux systèmes doivent réaliser des correspondances entre des entrées et des sorties d'une part, et entre un signal d'erreur (la différence entre l'état attendu et l'état réel) et un signal de correction d'autre part. Or, comme le soulignent Ingvaldsen et Whiting (1997), à la différence d'un dispositif électronique pour lequel le concepteur peut modifier le dispositif, changer des paramètres et en observer les effets pour éventuellement réviser le mécanisme de rétroaction, il n'a pas été possible à ce jour d'apporter une preuve directe de l'existence des modèles internes (Faisal et al., 2008; Glimcher, 2003; Ingvaldsen & Whiting, 1997; Karniel, 2011; Körding, 2007). En effet, même si des hypothèses sur la localisation cérébrale de tels systèmes de représentations sont privilégiées (dans leur revue de questions, Wolpert et al., 2001, et Karniel, 2011, évoquent notamment le cervelet), elles n'ont pas encore été confirmées, et ce malgré la modernisation des techniques d'investigation et le nombre d'études qui y sont consacrées.

Dès lors, un discrédit semble jeté sur ces approches computationnelles de l'apprentissage en tant que théories explicatives (Baer & Rosales-Ruiz, 2003; Donahoe, 2003; Gewirtz & Pelaez, 1996; Pelaez, Gewirtz, & Wong, 2008), dans la mesure où elles sont difficilement falsifiables. Comme le soulignait déjà Bachelard (1938) dans son exposé des « obstacles épistémologiques » (c'est-à-dire des causes de lenteur voire de stagnation dans la recherche scientifique), une explication ne peut se réduire à une description, aussi précise soit-elle, ni à une substantialisation, soit le fait de transformer une abstraction en une entité matérielle, concrète – particulièrement visible dans le vocabulaire employé : un modèle interne est-il une structure interne, ou sa formalisation ? La confusion n'est pas souvent levée par les auteurs.

Par ailleurs, si l'on concédait l'existence et le mode de fonctionnement de tels modèles internes – ainsi que leur capacité à guider le comportement – d'autres questions fondamentales se posent. Parmi elles, celles qui portent sur l'apprentissage sont particulièrement intrigantes.

Les théories computationnelles conceptualisent l'adaptation comme des changements dans les paramètres du système, qui auraient lieu « dans le but d'améliorer la performance ultérieure de ce système », et ce dans des tâches bien définies (telle l'adaptation saccadique). Selon ces théories, l'apprentissage, quant à lui, correspondrait à des changements structurels qui auraient des implications sur les comportements, comme l'adoption de stratégies optimales (telle la *main sequence*), et qui permettraient au système d'optimiser sa gestion des différents coûts associés aux mouvements – comme le compromis entre vitesse et précision (Karniel, 2011; Körding, 2007; Wolpert et al., 2001).

Plusieurs interrogations peuvent être soulevées. D'une part, d'où proviendrait la référence, l'idéal défini *a priori*, que le système utiliserait pour générer et éventuellement corriger ses actions de manière si optimale ? Quels que soient le sens et la réalité que l'on prête au concept de mémoire, nécessairement impliquée dans cette question, il reste que le système biologique étudié (en l'occurrence le système oculo-moteur) est alors supposé capable de réaliser une abstraction et de détenir quelque part la représentation de ce mouvement idéal, afin de pouvoir guider ses actions. Les théories computationnelles n'apportent pas de réponse à cet autre problème que pointent aussi Ingvaldsen et Whiting (1997).

D'autre part, l'implication de mécanismes adaptatifs distincts et hautement spécialisés – par exemple, certains auteurs, comme Wolpert et Kawato (1998), postulent l'existence de modèles internes multiples pour rendre compte de l'apprentissage de tâches diverses dans des contextes variés – peut elle aussi déconcerter, au regard non seulement du principe de parcimonie, mais aussi des ressources cognitives que cela pourrait représenter. Par ailleurs, les théoriciens de l'approche computationnelle semblent reconnaître eux-mêmes que la distinction entre adaptation et apprentissage n'est pas encore clarifiée (Karniel, 2011) : pourquoi en effet envisager des mécanismes adaptatifs différents ?

En proposant dans le cadre de cette thèse une approche opérante du contrôle des saccades, il s'agit d'expliquer l'apprentissage oculo-moteur par un mécanisme général d'apprentissage – un mécanisme qui ne se restreigne pas à l'apprentissage de comportements oculo-moteurs ni même sensori-moteurs.

II. L'apprentissage oculo-moteur selon l'analyse expérimentale du comportement : un mécanisme général d'apprentissage

L'analyse expérimentale du comportement définit l'apprentissage comme la modification durable des comportements en fonction des interactions avec l'environnement. Cette formulation est assez générale pour être compatible avec celle que donnent les théories classiques du contrôle moteur. Toutefois, ces approches s'opposent quant à la localisation des paramètres qui contrôlent les modifications comportementales.

Pour les théories classiques du contrôle moteur, l'apprentissage mobiliserait des modèles internes impliqués dans des calculs de gains, ou encore des optimisations de fonctions de coûts. En revanche, l'analyse expérimentale du comportement, héritée du positivisme logique et du behaviorisme méthodologique de Watson, se centre sur l'analyse des comportements observables (quels qu'ils soient), et refuse de tirer des inférences à propos des mécanismes internes qui contrôleraient le comportement. Au delà d'une différence méthodologique, l'analyse du comportement se distingue des théories computationnelles avant tout dans la mesure où elle confère aux conséquences des comportements un rôle central dans l'apparition, le maintien, et la modification de la forme de ces comportements eux-mêmes.

2.1. Les principes fondamentaux de l'apprentissage opérant

Selon l'analyse expérimentale du comportement, le rôle de ces conséquences est même *déterminant* dans l'apprentissage. En d'autres termes, cette théorie avance que le comportement est contrôlé par ses propres conséquences (ex : Skinner, 1953).

L'apprentissage serait l'équivalent au niveau comportemental de la sélection naturelle au niveau des espèces (Donahoe, 2003; Donahoe et al., 1993; Skinner, 1981). En effet, la sélection naturelle affecte les propriétés morphologiques et physiologiques des individus : si un organisme possède des caractéristiques qui lui sont bénéfiques, alors il aura davantage de chances de se reproduire et de transmettre ses gènes à sa descendance. L'analyse du comportement postule que l'apprentissage relèverait d'un mécanisme similaire de sélection au niveau des comportements :

ceux-ci ont davantage de chances d'être émis quand par le passé ils ont été suivis de conséquences bénéfiques pour l'organisme que lorsqu'ils ont été suivis d'aucune conséquence ou d'une conséquence aversive. Dès lors, c'est en observant les conséquences passées d'un comportement qu'il peut être possible d'en prédire les occurrences futures. En outre, c'est en manipulant les conséquences actuelles qu'il devrait être possible de contrôler la probabilité de ces occurrences (Skinner, 1953).

Le terme de renforcement désigne les procédures, les opérations, qui consistent à manipuler les conséquences d'un comportement dans le but de faire augmenter sa probabilité d'apparition (ou de maintenir cette probabilité élevée). Dans le cas d'un renforcement positif, un stimulus appétitif (bénéfique pour l'organisme) est présenté à l'issue d'un comportement, tandis que le renforcement négatif consiste à retrancher un stimulus aversif. Par exemple, dans un jeu vidéo, atteindre certaines cibles est souvent renforcé positivement par le gain de points, les comportements d'atteinte de ces cibles (comme se diriger dans certaines zones de l'écran plutôt que d'autres) voient leur probabilité s'accroître, et s'éloigner des zones peuplées par un adversaire est renforcé négativement – le joueur peut éviter alors un coup fatal.

Le terme de punition correspond aux procédures qui visent à faire diminuer la probabilité d'occurrence d'un comportement, soit en présentant un stimulus aversif à l'issue du comportement dans le cas d'une punition positive, soit en retranchant un stimulus appétitif dans le cas d'une punition négative. Dans l'exemple du jeu vidéo, si le joueur s'aventurant trop près du camp adverse est atteint par l'ennemi, perdant ainsi des points, son comportement est puni négativement.

On parle d'extinction lorsque l'on cesse de présenter la conséquence d'un comportement qui auparavant contrôlait sa fréquence d'apparition. Ce terme réfère également à la diminution de la fréquence de ce comportement souvent observée dans ce cas.

Quant à la contingence, elle désigne la relation entre la probabilité d'émettre un comportement et la probabilité d'apparition des stimulus dans l'environnement.

Le comportement est défini comme une classe de réponses, qui peuvent avoir la même forme ou des formes différentes, mais qui sont avant tout déterminées par la même *fonction* : un comportement *opérant* particulier correspond donc à l'ensemble des réponses qui font apparaître ou qui retranchent les mêmes éléments de l'environnement.

Quant à l'agent renforçateur, il correspond à cette conséquence fonctionnelle, à cet événement qui suit l'émission d'un comportement et qui détermine la probabilité d'émission des

réponses. Cette définition de l'agent renforçateur serait circulaire si elle n'était pas établie a posteriori, c'est-à-dire après la mise en évidence – expérimentale – de la capacité effective de cette conséquence à agir sur le comportement. De même, c'est a posteriori que l'on peut déterminer la valeur d'un agent renforçateur : en *fonction* de ses effets sur le comportement (Francisco, Borrero, & Sy, 2008; Hodos & Kalman, 1963; Skinner, 1953).

L'analyse du comportement consiste donc à manipuler systématiquement les éléments de l'environnement, en vue de déterminer expérimentalement leurs relations fonctionnelles avec le comportement étudié (le comportement étant fonction des événements de l'environnement, et les conséquences apparaissant elles-mêmes en fonction des comportements).

En ce qui concerne les mouvements oculaires, leurs conséquences immédiates sont de nature visuelle – il s'agit de la vision nette (ou non) de la cible. Ce serait donc ces conséquences qui contrôleraient, maintiendraient et sélectionneraient les mouvements oculaires, qui détermineraient leur morphogenèse.

Conférer un rôle fonctionnel à la vision nette de la cible à l'issue de la saccade implique une conception du signal à l'origine de l'adaptation saccadique différente de ce que postulent les théories classiques du contrôle moteur. D'une part, la vision nette de la cible ne peut être définie comme un agent renforçateur indépendamment de sa capacité à contrôler le comportement – ce qui n'est pas le cas du signal impliqué dans les calculs et prédictions postulés par les théories classiques. D'autre part, la conséquence fonctionnelle de la saccade n'a pas de signe, elle n'est ni positive ni négative.

Dans cette seconde partie sont présentés les principaux résultats expérimentaux sur lesquels s'appuie cette hypothèse de la vision nette de la cible fonctionnant comme un agent renforçateur.

2.2. La « sélection de la cible » sous l'angle de la théorie du renforcement opérant

Les études sur les mouvements oculaires réalisées en laboratoire permettent de contrôler les conséquences des saccades (les expériences employant le paradigme de double saut en sont de bons exemples).

Les premières recherches sur le renforcement opérant des mouvements oculaires chez l'être humain ont consisté à manipuler la probabilité d'apparition de signaux contingents aux saccades.

Elles avaient pour but de tester les effets de différents programmes de renforcement sur ces comportements (Rosenberger, 1973; Schroeder & Holland, 1968, 1969).

En effet, les études en analyse expérimentale du comportement avaient alors mis en évidence le fait que les taux de réponses (comme des coups de bec sur des clés chez le pigeon) variaient différemment selon les programmes de renforcement mis en place. Ces programmes peuvent être par exemple des programmes de renforcement à ratio, dans lesquels le renforcement dépend du nombre de réponses émises (indépendamment du moment où elles sont émises). Leurs effets sur les taux de réponses se distinguent de ceux produits par les programmes de renforcement à intervalle, dans lesquels les réponses ne sont renforcées que si elles sont émises après un intervalle de temps prédéfini (ex : Skinner, 1953).

Dans la recherche de Schroeder et Holland (1968), les participants sont placés devant un dispositif électronique comportant plusieurs pointeurs (de petites aiguilles qui peuvent changer d'orientation). Ils doivent en rapporter les changements de direction. Les auteurs mettent en place des programmes de renforcement multiples. Un programme de renforcement multiple consiste à alterner plusieurs composantes, celles-ci correspondant à des programmes de renforcement indépendants pour un seul et même comportement (en l'occurrence, le fait de déplacer le regard d'un pointeur vers un autre, quel qu'il soit). Chaque composante est signalée par un stimulus antécédent spécifique. Rosenberger (1973) et Schroeder et Holland (1969) utilisent le même dispositif, mais ils appliquent des programmes de renforcement concurrents. C'est-à-dire que plusieurs programmes de renforcement opèrent de manière indépendante et simultanée pour différents comportements (un comportement étant alors défini par la cible regardée).

Les résultats montrent que la fréquence des saccades dépend effectivement du taux programmé d'apparition des signaux, et les patterns de réponses sont similaires à ce qui est observé pour d'autres comportements – comme les coups de bec des pigeons sur des clés. Les auteurs en concluent que les saccades oculaires sont sous le contrôle des contingences de renforcement.

Ces données sont corroborées par des recherches dans lesquelles des stimulus non visuels, comme de l'argent pour les humains ou du jus de fruits pour les singes, sont dispensés suite à des saccades précises, selon différents programmes de renforcement.

Par exemple, dans son étude de 1968 chez le singe, Berger met en place un programme de renforcement multiple des saccades supérieures à 10° (quelle que soit leur direction). Il obtient des résultats conformes à ceux de Schroeder et Holland (1968).

Selon les expériences, plusieurs paramètres peuvent être manipulés, comme la grandeur de la conséquence associée à chaque cible et/ou la probabilité de son apparition. Ainsi, une relation linéaire a été mise en évidence entre la proportion de saccades dirigées vers une cible d'une part et la quantité relative de jus de fruits (Platt & Glimcher, 1999) ou la probabilité relative de renforcement (Sugrue, Corrado, & Newsome, 2004) associées à cette cible d'autre part.

Les recherches de Stritzke, Trommershäuser et Gegenfurtner (2009) et Schütz et Gegenfurtner (2010) apportent un argument supplémentaire en faveur de l'influence des conséquences des saccades sur leur orientation précise à l'intérieur d'une cible. Dans leurs expériences chez l'être humain, une pénalité (financière) est associée à un distracteur situé non loin de cette cible (des items de 2° de diamètre), alors que cette dernière est associée à une récompense.

Ils observent non seulement que la très grande majorité, sinon la totalité, des saccades sont dirigées vers la cible (Stritzke et al., 2009), mais encore que les saccades réalisées vers la cible s'éloignent de la région proche de celle donnant lieu à une pénalité (Schütz & Gegenfurtner, 2010).

La théorie de l'apprentissage opérant définit le choix comme l'émission d'un comportement donné quand plusieurs comportements sont possibles. En l'occurrence, la « sélection de la cible » correspond donc à la réalisation d'une saccade particulière, vers une cible déterminée, plutôt que vers d'autres cibles.

Les recherches de Platt et Glimcher (1999), Schroeder et Holland (1969) et Sugrue et al. (2004) soulignent que la probabilité d'émission des saccades vers une cible précise suit la loi du *matching*. Cette loi formalise les situations de choix lorsque l'organisme est confronté à des programmes de renforcement concurrents (Herrnstein, 1961; Poling, Edwards, Weeden, & Foster, 2011). Elle prédit notamment que la probabilité d'émettre un comportement opérant donné, relativement aux autres comportements possibles, dépend de l'importance relative du renforcement programmé pour ce comportement.

Plus précisément, ces études de Platt et Glimcher (1999), Schroeder et Holland (1969) et Sugrue et al. (2004) chez le singe ainsi que celles de Stritzke et al. (2009) et Schütz & Gegenfurtner (2010) chez l'humain mettent en exergue le contrôle que peuvent prendre des stimulus discriminatifs sur l'émission des saccades oculaires.

Un stimulus discriminatif précède une réponse et signale la disponibilité d'un agent renforçateur. Cette association avec l'agent renforçateur résulte d'un apprentissage antérieur, plus

exactement d'un processus de *renforcement différentiel*. Ce dernier correspond au renforcement de certains comportements dans des situations précises, c'est-à-dire en présence des stimulus qui deviendront ensuite des stimulus discriminatifs, tandis qu'en leur absence ces comportements ne sont pas renforcés. En vertu de cette association, la présence d'un stimulus discriminatif change la probabilité d'émission des comportements qui dans le passé ont été suivis par ces agents renforçateurs.

En d'autres termes, c'est parce qu'un organisme exécute des comportements différemment selon la présence ou l'absence de tels stimulus que ceux-ci sont appelés stimulus discriminatifs ; on parle aussi de *contrôle discriminatif*. Quant à la contingence à trois termes, elle fait référence à la relation entre un comportement, les stimulus discriminatifs et les conséquences de ce comportement.

Les études évoquées ci-dessus contribuent donc à montrer que l'orientation des saccades oculaires peut être placée sous un tel contrôle. Cela implique que l'organisme a appris une association entre une cible visuelle, fonctionnant comme un stimulus discriminatif, et la conséquence des saccades – le jus de fruits, de l'argent, la détection d'un changement de direction d'un pointeur. Ces recherches suggèrent qu'en dehors du laboratoire, l'orientation des saccades dépendrait aussi de l'apprentissage d'associations entre des cibles et des agents renforçateurs, à savoir des signaux visuels qui permettent d'interagir avec l'environnement.

2.3. Les saccades oculaires possèdent des dimensions opérantes

On peut faire l'hypothèse que les contingences de renforcement à trois termes ne contrôlent pas uniquement la direction des saccades. En effet, les conséquences de ces dernières, en renforçant certains mouvements, pourraient aussi sélectionner certaines de leurs propriétés.

Les propriétés d'un comportement sont appelées des *dimensions opérantes* lorsque ce sont elles qui déterminent l'apparition de la conséquence. Dès lors, cette conséquence devrait contrôler aussi les valeurs que vont prendre ces dimensions – ce qui demande à être prouvé expérimentalement (dans l'exemple du jeu vidéo, si le joueur ne marque des points que lorsqu'il accède à certaines zones rapidement, avant son adversaire, alors les comportements d'atteinte dont la latence est courte et la vitesse élevée ont une probabilité d'apparition plus importante que les mouvements lents aux latences longues).

2.3.1. Le contrôle discriminatif de la latence et du pic de vitesse des saccades : recherches chez le singe

Chez le singe, le paradigme dit de « *one direction-rewarded task* », qui a pour principal objectif d'investiguer les corrélats neuronaux des changements induits par le renforcement des saccades, permet aussi, au niveau comportemental, de mettre en évidence le contrôle de la latence et du pic de vitesse par les contingences de renforcement (voir Hikosaka, Nakamura, & Nakahara, 2006, pour une revue de littérature).

Dans cette procédure, illustrée par la figure 6, les animaux doivent émettre une saccade vers une cible apparaissant aléatoirement à droite ou à gauche d'un point de fixation. Les saccades exécutées dans une direction donnée donnent lieu à l'obtention de jus de fruits. Quant aux saccades exécutées dans l'autre direction, elles doivent être précises (c'est-à-dire atteindre la cible) pour que les essais suivants soient présentés – et finalement permettre au singe d'obtenir l'agent renforçateur. Après un certain nombre d'essais, la direction associée à l'arrivée du liquide est changée. Les chercheurs comparent les saccades émises dans la direction associée à l'obtention du jus de fruits à celles qui n'en sont pas suivies.

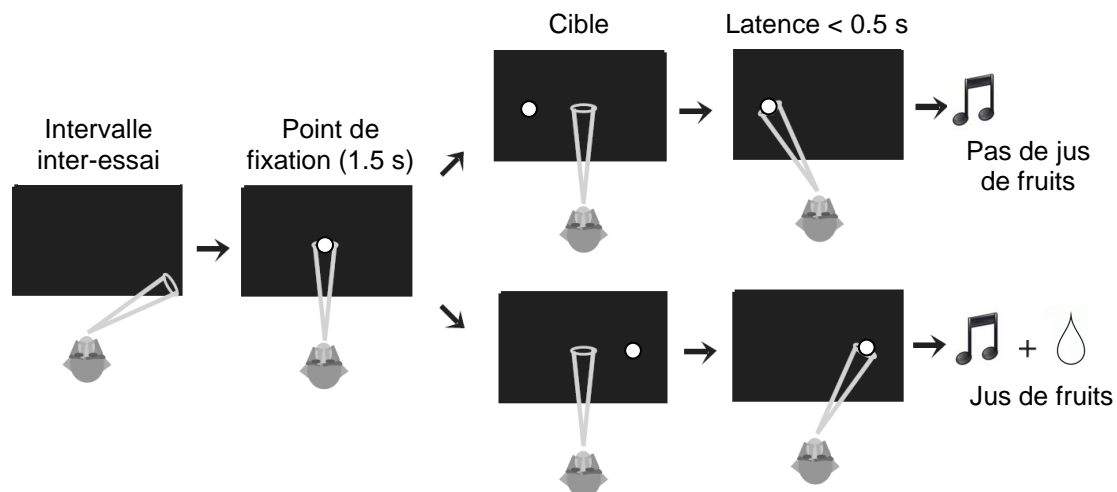


Figure 6. Procédure expérimentale de la « *one direction-rewarded task* ». Dans cette condition expérimentale, les saccades dirigées vers la droite sont renforcées avec du jus de fruits. Dans la suivante, l'arrivée de jus de fruits suit les saccades dirigées vers la gauche ; ces conditions sont alternées plusieurs fois. (source : Lauwereyns et al., 2002.)

En utilisant ce paradigme, Takikawa, Kawagoe, Itoh, Nakahara et Hikosaka (2002) constatent que les saccades émises dans la direction donnant lieu à un agent renforçateur alimentaire ont des trajectoires plus rectilignes et des pics de vitesse plus élevés que les autres. Les latences de ces saccades suivies de l'obtention de liquide sont aussi plus courtes (Lauwereyns, Watanabe, Coe, & Hikosaka, 2002; Nakamura & Hikosaka, 2006; Takikawa et al., 2002; Watanabe, Lauwereyns, & Hikosaka, 2003).

En outre, ces recherches indiquent que quand la direction annonçant l'obtention de jus de fruits est changée, un seul essai suffit pour réduire les latences des saccades nouvellement récompensées avec le liquide. Cela révèle que la direction, en signalant l'arrivée de cette récompense, fonctionne comme un stimulus discriminatif et contrôle les propriétés des saccades.

Notons que ce contrôle est effectif même si l'agent renforçateur n'est pas contingent à ces propriétés (c'est-à-dire qu'il n'en dépend pas) : en effet, son obtention est déterminée uniquement par la direction de la saccade. Hypothétiquement, les changements immédiats des latences et des pics de vitesses pourraient être le reflet d'un apprentissage antérieur, dans lequel un processus de renforcement, contingent à ces propriétés, aurait augmenté la probabilité d'émission des saccades aux latences courtes et pics de vitesses élevés.

Ces observations sont corroborées par la recherche de Sohn et Lee (2006) qui porte sur des saccades émises en direction d'une cible qui se déplace en huit sur un damier, changeant de case après chaque saccade. La récompense est toujours associée à la même localisation. Sohn et Lee (2006) observent que la latence des saccades est corrélée positivement avec le nombre de saccades restant avant l'arrivée à ladite localisation, ainsi qu'avec la distance qui sépare la cible de la saccade à émettre et cette localisation. C'est cette dernière localisation qui, annonçant l'arrivée de la récompense, suscite la réduction de la latence de ces saccades, vraisemblablement en raison d'un apprentissage antérieur de contingences de renforcement portant sur cette propriété.

2.3.2. Recherches chez l'être humain

Des études réalisées chez l'être humain indiquent également que la latence et le pic de vitesse des saccades sont des propriétés contrôlées par des contingences de renforcement à trois termes.

Par exemple, Montagnini et Chelazzi (2005) demandent à leurs participants de réaliser une saccade vers une cible qui est remplacée soit par la lettre E, soit par un E inversé après un délai très bref (correspondant à la somme de la latence médiane et de la durée médiane des saccades). Les participants doivent discriminer ce stimulus sachant qu'il n'est illuminé que 13 ms. Par

conséquent, si la saccade n'est pas terminée assez tôt, les chances de réussir la tâche diminuent. Montagnini et Chelazzi (2005) observent effectivement une réduction significative de la latence moyenne (de 15,60%) par rapport à une condition contrôle n'impliquant aucune tâche de choix forcé.

Les résultats peuvent être interprétés sous l'angle de l'analyse du comportement, même s'il ne s'agit pas d'une expérience d'apprentissage par renforcement opérant à proprement parler. En effet, dans cette étude, c'est plutôt le contexte qui serait manipulé. Contrairement à la condition contrôle, dans le contexte de la condition expérimentale, le point de fixation fonctionne comme un stimulus discriminatif : il signale l'arrivée rapide d'une cible qui ne sera plus disponible si la latence est trop longue. Puisqu'en présence de ce contexte précis, les latences sont courtes, il est possible d'en déduire que la latence des saccades est placée sous le contrôle du stimulus – autrement dit que la latence dépend des contingences de renforcement ayant probablement été apprises avant l'expérience.

Par ailleurs, Montagnini et Chelazzi (2005) constatent aussi une augmentation significative du pic de vitesse, ce qui laisse penser que la vitesse pourrait être une autre dimension opérante des saccades oculaires.

L'influence des conséquences des mouvements sur le pic de vitesse fait également l'objet d'une étude réalisée par Xu-Wilson, Zee et Shadmehr (2009). Pendant que leurs participants regardent un point de fixation au début d'un essai, les auteurs présentent une image (par exemple des visages humains ou des pixels sans signification) en vision périphérique. Ces images varient aléatoirement. Puis la cible apparaît, les participants doivent alors faire une saccade dans sa direction et 300 ms après la fin du mouvement, l'image vue précédemment est présentée en guise de conséquence. Xu-Wilson et al. (2009) observent une augmentation significative (de 5.5 °/sec) du pic de vitesse moyen des saccades réalisées avant le visage humain par rapport à celui des saccades réalisées avant la présentation des pixels. De même, la durée des premières saccades est significativement supérieure (de 1.7 ms) à celle des secondes. Les auteurs considèrent que ces propriétés dépendent de la valeur de la conséquence préalablement annoncée.

Puisque, dans cette expérience également, les conséquences des saccades ne dépendent pas des pics de vitesse, on peut supposer, selon la même logique que précédemment, que les changements de cette propriété relèveraient ici aussi d'un contrôle par des stimulus discriminatifs – par les images qui apparaissent en vision périphérique. Cela expliquerait que les différences soient visibles sur un seul essai.

En ce qui concerne l'amplitude des saccades, Montagnini et Chelazzi (2005), Xu-Wilson et al. (2009), de même que Takikawa et al. (2002) dans leurs études chez le singe, soulignent qu'elle n'est pas affectée, ce qui du reste est imposé par les procédures employées. En effet, l'agent renforçateur n'est contingent qu'à une propriété : la direction *précise* des saccades – cette précision est nécessaire pour discriminer les items dans la recherche de Montagnini & Chelazzi (2005), et quant aux saccades des singes de Takikawa et al. (2002), si elles amènent leur fovéa trop loin de l'image de la cible, l'essai, considéré comme incorrect, est réitéré.

En revanche, le fait que le pic de vitesse et la durée soient modifiées par les conséquences des saccades semble indiquer que dans ces conditions, la *main sequence*, cette relation prétendument stéréotypée entre amplitude et vitesse ou durée des saccades, pourrait être altérée et placée sous le contrôle des contingences de renforcement.

2.3.3. La manipulation des propriétés saccadiques par des procédures de renforcement

Toutefois, dans les études ci-dessus, cette hypothèse du contrôle des propriétés saccadiques telles que la latence et le pic de vitesse par leurs propres conséquences n'a pas été testée directement. Une autre interprétation des résultats en termes d'arousal pourrait être proposée. En effet, la simple présentation de stimulus dont l'intensité physique est importante ou de stimulus qui peuvent fonctionner comme des agents renforçateurs peut entraîner une hausse du niveau d'activité générale de l'organisme. L'apparition de ces stimulus n'est alors pas nécessairement contingente aux comportements. Cela se traduit, au niveau neuronal comme comportemental, par une augmentation transitoire de l'intensité des réponses, qui dépend de la fréquence de la présentation ainsi que de l'intensité de ces stimulus (Killeen, Hanson, & Osborne, 1978; Schultz, 2010). Dans la mesure où aucune diminution des latences ni des pics de vitesse n'a été démontrée dans ces expériences de « *one direction-rewarded task* », ni dans celle de Montagnini et Chelazzi (2005) ou de Xu-Wilson et al. (2009), l'hypothèse d'un effet d'arousal ne peut être écartée.

Des études récentes réalisées dans notre laboratoire montrent que le contrôle de la latence et du pic de vitesse des saccades oculaires peut effectivement relever d'un apprentissage opérant. Ainsi, Madelain, Champrenaut et Chauvin (2007) montrent qu'il est possible de modifier systématiquement la distribution des temps de réaction en renforçant directement les latences des

saccades. C'est-à-dire que les conséquences des saccades apparaissent alors en *fonction* de cette propriété.

Les participants de l'étude de Madelain et al. (2007) doivent effectuer une tâche de choix forcé. Après l'extinction d'un point de fixation au centre d'un écran, deux cibles apparaissent en vision périphérique. Ils doivent réaliser une saccade vers la cible la plus lumineuse et les réponses correctes sont renforcées par un son selon plusieurs critères de latence, dans quatre conditions expérimentales différentes. Dans les deux premières conditions, la médiane des latences doit rester constante et la dispersion doit être augmentée ou bien réduite. Dans les deux autres conditions ces deux paramètres doivent être modifiés dans la même direction (augmentés ou diminués). Les auteurs obtiennent des changements de distributions qui dépendent effectivement des contingences de renforcement. Ils parviennent notamment à manipuler indépendamment la médiane et la variance des latences saccadiques, montrant ainsi que cette propriété des mouvements oculaires dépend d'un apprentissage opérant.

En ce qui concerne le pic de vitesse, une étude préliminaire indique que cette propriété peut elle aussi être contrôlée par ses propres conséquences (Madelain & Paeye, 2011). Dans une première condition expérimentale, un son signalant le gain de points est émis après les saccades dont les pics de vitesse sont les plus élevés. Puis, dans la condition suivante, ce son est émis à la suite des saccades dont les pics de vitesse sont les plus faibles. Selon les conditions, une augmentation de 14% et une diminution de 12% sont respectivement observées, ce qui conforte l'hypothèse du contrôle du pic de vitesse saccadique par les conséquences des mouvements.

Ces recherches confirment que latence et pic de vitesse sont des dimensions opérantes, dans la mesure où ces propriétés sont directement contrôlables par la manipulation de leurs propres conséquences – elles peuvent être expérimentalement augmentées *et* diminuées par rapport à la ligne de base.

2.3.4. Les corrélats neuronaux des changements induits par le renforcement des saccades

Le paradigme de la « *one direction-rewarded task* » permet de mettre en évidence des modifications d'activation neuronale corrélées avec les changements des propriétés saccadiques induits par le renforcement (Hikosaka et al., 2006). Ces modifications d'activation ont lieu non seulement après, mais aussi avant que la direction de la saccade soit indiquée. Cela signifie que le contexte (en l'occurrence l'association entre la récompense et une direction donnée, association

qui reste la même pendant toute une session expérimentale) est corrélé avec des patterns d'activité neuronale particuliers – tout comme les stimulus discriminatifs peuvent être corrélés avec certaines propriétés saccadiques.

Les recherches qui emploient ce paradigme ont principalement porté sur les ganglions de la base. Ainsi, Lauwereyns et al. (2002) et Watanabe et al. (2003) observent que l'activité anticipatrice des neurones situés dans le noyau caudé change systématiquement en fonction de la direction associée à la récompense. Des changements sélectifs d'activité neuronale en fonction de la direction associée à la récompense sont également relevés en aval du noyau caudé : dans la substance noire (Sato & Hikosaka, 2002), et dans le colliculus supérieur (Ikeda & Hikosaka, 2003).

En outre, les expériences de choix dans lesquelles les singes doivent faire une saccade vers une cible associée à un certain agent renforçateur révèlent des modifications d'activité neuronale dans l'aire intrapariétale latérale (Glimcher, 2003; Trommershäuser, Glimcher, & Gegenfurtner, 2009). Cette activité, corrélée avec les comportements de choix, y est modulée par la quantité relative de jus de fruits obtenue ou bien par la probabilité qu'une réponse particulière soit suivie d'une récompense (Platt & Glimcher, 1999; Sugrue et al., 2004).

Les mécanismes de la modulation du système saccadique par le renforcement ne se résument en aucun cas à ces changements localisés (ex. : Glimcher, 2003; Hikosaka et al., 2006; Trommershäuser et al., 2009) – d'autres changements peuvent avoir lieu, mais ne pas avoir fait l'objet de telles investigations. La dopamine semble aussi particulièrement impliquée dans l'apprentissage des saccades oculaires.

En premier lieu, les neurones à dopamine se projettent massivement dans les ganglions de la base (dans le noyau caudé notamment) et dans le cortex dorsolatéral, des structures engagées dans le contrôle des saccades oculaires, ce qui semble plaider en faveur de cette hypothèse (Hikosaka et al., 2006).

En second lieu, les résultats de Nakamura et Hikosaka (2006) suggèrent que le système dopaminergique joue un rôle important dans les effets du renforcement sur les saccades. Ils injectent des antagonistes de la dopamine dans le noyau caudé de leurs singes et observent que la différence de temps de réaction entre les saccades renforcées et non renforcées par du jus de fruits est amoindrie, par rapport à des situations contrôles sans injection.

Quant à Bayer & Glimcher (2005), ils entraînent des singes à faire des saccades qui ne sont renforcées par du liquide que si leur latence se situe dans un intervalle déterminé (par exemple,

entre deux et trois secondes). Cet intervalle est divisé en cinq, chaque subdivision donnant lieu à une quantité de liquide qui augmente graduellement. Cet intervalle change tous les 100 essais environ. Les auteurs peuvent donc tirer profit des modifications de contingences de renforcement pour observer l'activité des neurones à dopamine : lorsque le singe obtient du liquide, l'activité de ces neurones dépend de la différence entre la récompense actuelle et celle qu'il a obtenue lors des essais antérieurs.

Ces corrélats neuronaux du renforcement des saccades confortent l'idée que ces dernières constituent un comportement opérant.

En effet, les données obtenues dans les recherches sur les saccades oculaires sont cohérentes avec celles qui ont été réalisées sur l'apprentissage d'autres comportements (ex : Hollerman, Tremblay, & Schultz, 1998), tels que des tâches familières ou des comportements liés à la consommation de drogue (voir aussi : Schultz, 2006, 2010, pour des revues de littérature) : ces recherches mettent aussi en évidence que les temps de réaction sont plus courts quand une récompense est annoncée que quand elle ne l'est pas, et que les performances comportementales s'accompagnent de changements d'activité neuronale dans les ganglions de la base.

Ces études ont également souligné le rôle prépondérant que joue la dopamine dans l'apprentissage des comportements (Schultz, 2006, 2010), même si les mécanismes dopaminergiques ne sont pas totalement élucidés aujourd'hui. Elles révèlent notamment que l'activité des neurones dopaminergiques augmente quand un événement appétitif apparaît alors que sa probabilité d'occurrence était faible. Inversement, l'activité neuronale est inhibée quand cet événement appétitif n'apparaît pas alors que sa probabilité d'occurrence était forte. Par ailleurs, l'activité des neurones ne change pas si cet événement apparaît conformément à ce qui se produisait dans le passé.

Cette corrélation de l'activité des neurones dopaminergiques avec les différences entre la situation actuelle (par exemple, la récompense obtenue suite à un comportement) et les contingences antérieures amène des auteurs comme Schultz et ses collègues à parler « d'erreurs de prédiction », qui correspondraient aux signaux guidant l'apprentissage (Montague, Hyman, & Cohen, 2004; Schultz, 2000, 2006, 2010). Ainsi, selon eux, la *contiguïté* (la proximité temporelle entre les événements) et la *contingence* (la probabilité d'apparition de la récompense dans un contexte donné) seules ne suffisent pas pour rendre compte de la modification des comportements. En effet, il faut aussi que « l'apparition d'une récompense soit surprenante ou imprévue pour qu'une association avec un stimulus ou une action soient apprises » (Schultz,

2000, p. 200, traduction par nos soins), un principe formalisé auparavant par Rescorla et Wagner (1972) notamment.

Au niveau comportemental, le parallèle peut être fait avec la notion de « divergence comportementale » – « *behavioral discrepancy* » en anglais (Guerra & Silva, 2010). Ce concept a été introduit par Donahoe et Palmer qui proposent un « principe de renforcement unifié » (« *unified principle of conditioning* »), c'est-à-dire un principe qui rend compte de la sélection comportementale où les conditionnements répondant et opérant interviennent simultanément (Donahoe, 2003; Donahoe et al., 1993; Donahoe & Palmer, 1994).

La divergence comportementale désigne le moment où un changement dans les actions en cours se produit suite à l'apparition d'un stimulus intense ou susceptible de fonctionner comme un agent renforçateur (par exemple l'arrivée de nourriture dans la mangeoire pour un rat dans une expérience).

Plus précisément, Donahoe et Palmer suggèrent que cet événement entraîne, par des associations de type répondant, l'émission de certaines réponses particulières (comme la salivation), spécifiques à ce stimulus – alors que l'organisme était en train d'émettre d'autres comportements sans lien avec ce stimulus. Par conditionnement répondant, les stimulus qui précèdent immédiatement l'arrivée de ce nouveau stimulus (inconditionnel d'un point de vue du conditionnement répondant) vont y être associés.

Dans le même temps, par conditionnement opérant, les comportements (comme appuyer sur un levier) qui précèdent ou sont concomitants à l'arrivée de ce nouveau stimulus (agent renforçateur putatif d'un point de vue opérant) vont également y être associés. Ce n'est qu'a posteriori que le stimulus qui a suivi le changement de comportement ne peut être considéré comme un agent renforçateur d'un point de vue opérant : à condition qu'il produise des modifications dans le comportement.

La divergence comportementale permet donc la sélection des associations entre les stimulus, les réponses et les agents renforçateurs, à condition que ces événements soient contigus dans le temps et que leurs occurrences soient corrélées. Pour qu'elle se produise, elle nécessite également la présence de variations dans le comportement comme dans l'environnement.

2.3. La variabilité comportementale, une dimension opérante

La conception de la variabilité comme essentielle à l'apprentissage se distingue de celle des théories computationnelles, qui la considèrent souvent comme le reflet d'écarts par rapport à une norme. Dans les modèles proposés par ces théories, la variabilité prend la forme des écarts-types de distributions statistiques résumant les données. Quant aux théories du contrôle optimal, elles la conçoivent comme les conséquences délétères d'un bruit neuronal aléatoire, incontrôlable, qui affecterait chaque étape des comportements sensori-moteurs (Faisal et al., 2008). Selon elles, la variabilité ferait partie des coûts que le système, adoptant des stratégies optimales, chercherait à minimiser (ex : Harris, 1995, Harris & Wolpert, 1998, 2006, van Beers, 2008).

S'il n'est pas question de nier les effets du bruit neuronal sur les variations comportementales, une approche opérante du contrôle moteur pourrait envisager cette propriété comme dépendant également des contraintes que l'environnement exercerait sur elle. On peut supposer par exemple que lors de la lecture, un faible niveau de variabilité soit profitable dans la mesure où l'orientation de la fovéa vers des localisations spécifiques peut donner lieu à la perception distincte de chaque mot. En revanche, lorsque l'observateur scrute une page blanche, les mouvements oculaires ne sont probablement pas aussi stéréotypés. Il est possible que dans ce cas la variabilité ne soit pas une dimension pertinente puisque les conséquences des mouvements sont les mêmes quelle que soit leur orientation. La variabilité traduirait alors l'influence de contraintes environnementales moins strictes au regard de cette propriété.

Autrement dit, la variabilité ferait partie des dimensions opérantes des saccades, ces propriétés qui à la fois déterminent l'apparition des agents renforçateurs et sont contrôlées par ces derniers, au même titre que d'autres propriétés comme la vitesse, l'orientation ou la latence. On peut faire l'hypothèse que le contrôle de la variabilité oculo-motrice relèverait donc (en partie) du même mécanisme général d'apprentissage que celui qui détermine les autres dimensions comportementales.

L'analyse expérimentale du comportement n'ayant que peu investigué les comportements sensori-moteurs, cette hypothèse s'appuie sur un certain nombre de recherches qui portent sur la variabilité d'autres comportements, principalement celle de séquences d'actions (voir Lee, Sturme, & Fields, 2007; Neuringer, 2002, 2009, pour des revues de littérature). Ces études, qui attestent que l'une des caractéristiques d'un comportement opérant est que sa variabilité est modifiable par les contingences de renforcement, font l'objet de cette dernière section qui s'éloigne donc du domaine sensori-moteur.

2.3.1. Définition et investigation de la variabilité opérante

De manière générale, la variabilité caractérise les relations entre les membres d'une même classe – en l'occurrence, les réponses d'une classe comportementale donnée. Envisager la variabilité comme une dimension opérante signifie que le renforcement détermine non seulement la probabilité d'apparition d'un comportement donné, mais aussi la probabilité d'apparition de certaines *formes* de réponses (ex. : Boulanger, Ingebos, Lahak, Machado, & Richelle, 1987; Neuringer, 2002, 2009). En d'autres termes, le renforcement conditionne la probabilité que les dimensions (par exemple, la durée, la latence, la localisation, etc.) de ce comportement prennent certaines valeurs. Soit ces valeurs sont extrêmement proches les unes des autres, lorsque le comportement est très stéréotypé, soit ces valeurs sont différentes et c'est alors que le comportement est qualifié de variable.

Il existe plusieurs moyens de modifier le niveau de variabilité comportementale, par exemple en manipulant les caractéristiques des stimulus (voir l'étude de Eckstein, Beutter et Stone (2001) pour une illustration dans le domaine des mouvements oculaires : le contraste des cibles que les participants doivent chercher dans du bruit visuel a un effet sur la variabilité de l'orientation des premières saccades au début des essais). Par ailleurs, les caractéristiques des participants eux-mêmes peuvent avoir une influence sur la variabilité (ainsi elle peut être accrue par l'absorption de drogue ou par des désordres neurologiques). Il ne s'agit pas à proprement parler de variabilité opérante si cette propriété n'est pas contrôlée par ses propres conséquences.

Par ailleurs, la variabilité comportementale peut également dépendre des instructions (Hopkinson & Neuringer, 2003; Joyce & Chase, 1990; Souza, Pontes, & Abreu-Rodrigues, 2012) ou de règles que les participants se forgent (Maes, 2003; Maes & Van der Goot, 2006; Souza et al., 2012). Dans ce cas, on peut faire l'hypothèse que la variabilité observée reflèterait (en partie) des contingences apprises antérieurement (Lattal & Neef, 1996). Parfois, ces apprentissages antérieurs ne permettent pas de répondre de manière optimale aux exigences imposées par les contingences en vigueur, c'est-à-dire d'obtenir le plus d'agents renforçateurs possibles (Galizio, 1979).

En effet, le niveau de variabilité induit par ces antécédents verbaux peut parfois être trop élevé pour permettre aux individus de détecter les changements de contingence. C'est ce qui semble se produire dans l'expérience de Maes et Van der Goot (2006). Lorsque les auteurs renforcent la production de séquences en fonction de la variabilité et appliquent simultanément un programme de renforcement d'une séquence particulière, les participants se comportent de

manière variable mais la probabilité d'émission de cette séquence cible n'augmente pas (contrairement à un groupe contrôle pour qui seule la séquence cible est renforcée).

En revanche, le niveau de variabilité induit par ces règles peut parfois être trop faible, et c'est seulement grâce aux programmes de renforcement opérant que les individus peuvent répondre aux exigences portant sur la variabilité. Ainsi, Neuringer et ses collaborateurs, qui demandent à leurs participants de se comporter de manière variable (Hopkinson & Neuringer, 2003), aléatoire (Neuringer, 1986), voire chaotique (Neuringer & Voss, 1993), observent une augmentation significative de la variabilité après la mise en place de programmes de renforcement, sans lesquels cette augmentation n'aurait probablement pas lieu.

Par rapport à ces procédés, le renforcement direct de la variabilité semble pouvoir induire des niveaux de variabilité bien déterminés.

Les premières études dévolues au renforcement opérant de cette dimension montrent qu'il est possible d'accroître expérimentalement la variabilité des formes globales des comportements (soit la « topographie »), en renforçant les réponses si et seulement si leurs formes diffèrent de celles qui ont été émises auparavant – cette procédure porte le nom anglais de « *lag procedure* ». Ainsi, Pryor, Haag, et O'Reilly (1969) ont augmenté le nombre de déplacements différents émis par des marsouins (par exemple des sauts ou des pirouettes). Quelques années plus tard, Goetz et Baer (1973) ont employé cette procédure pour amener des enfants à diversifier leurs assemblages réalisés avec un jeu de construction ; puis quand le renforcement était contingent à la construction d'une seule forme, la variabilité diminuait. Quant à Lee, McComas et Jawor (2002), ils ont augmenté la variabilité des réponses verbales émises par des enfants autistes en renforçant chaque phrase à condition qu'elle diffère de la précédente.

La plupart des recherches sur la nature opérante de la variabilité comportementale se focalise plutôt sur une dimension particulière de ces comportements. Ainsi, Blough, (1966) montre qu'il est possible d'augmenter expérimentalement la variabilité des intervalles inter-réponses de coups de bec donnés par des pigeons. Il met en place une procédure dite de « renforcement des réponses les moins fréquentes » (« *least frequent procedure* »). Il s'agit d'abord de calculer la distribution des fréquences de n réponses précédentes, chaque nouvelle réponse remplaçant la plus ancienne (par exemple, Blough (1966) utilise une fenêtre de 150 essais et répartit les intervalles inter-réponses en 16 classes). Ensuite la fréquence relative de la réponse actuelle ou de la classe contenant cette réponse est située dans cette distribution. Si cette fréquence est la plus faible, un agent renforçateur peut être délivré. La distribution des fréquences des réponses devrait

donc devenir de plus en plus uniforme au fur et à mesure de l'apprentissage, signe d'une augmentation de la variabilité comportementale.

Par la suite, des variantes de procédures basées sur les fréquences relatives des réponses ont été utilisées. Celle de Denney et Neuringer (1998) diffère de celle de Blough (1966) par la détermination d'une fréquence relative seuil pour le renforcement (d'où le nom de « procédure de seuil », ou « *threshold procedure* », que Denney et Neuringer (1998) lui donnent). C'est-à-dire que la réponse émise ne donnera lieu à un agent renforçateur que si sa fréquence relative est située en dessous d'une certaine valeur. Le nombre de réponses susceptibles de donner lieu à un agent renforçateur est donc plus élevé que ce que permet la procédure de Blough (1966). En utilisant cette procédure chez l'être humain, Ross et Neuringer (2002) parviennent à manipuler, de manière soit indépendante, soit simultanée, la variabilité de plusieurs dimensions (à savoir l'aire, la forme et/ou la localisation spatiale de rectangles dessinés sur un écran). Ces deux procédures se distinguent des *lag procedures* car deux réponses identiques successives peuvent être renforcées.

La majorité des études sur la variabilité comportementale porte cependant sur l'ordre d'actions réalisées en séquences : coups de bec sur deux clés chez des pigeons, appuis sur des leviers chez des rats, séquences d'appuis sur les touches d'un clavier chez les humains (ex. : Neuringer, 2002, 2009; Souza et al., 2012).

2.3.2. Les changements de variabilité induits indirectement par la fréquence de renforcement

Avant de tirer des conclusions sur la variabilité opérante proprement dite, il est nécessaire de s'assurer qu'elle est effectivement *contrôlée* par ses propres conséquences. Mettre en place une procédure de renforcement contingent à la variabilité ne suffit pas toujours pour le démontrer. En effet, la probabilité de renforcement elle-même a un effet sur la variabilité comportementale, que le renforcement soit contingent à la variabilité ou non.

Tout d'abord, plusieurs recherches indiquent que l'extinction a pour corollaire une augmentation de la variabilité comportementale (Boulanger et al., 1987; Lee et al., 2007). Antonitis (1951) le démontre dans une étude réalisée chez des rats ayant appris à introduire le museau dans une ouverture horizontale de 50 cm de long. Des périodes de renforcement continu (consistant à renforcer chaque réponse) et des périodes d'extinction sont alternées. La variabilité de la localisation des réponses diminue dans le premier cas, même si le renforcement n'est pas

contingent à cette propriété. En revanche, elle augmente lors de l'extinction – pendant laquelle une augmentation transitoire du taux de réponses a d'abord lieu. Plus récemment, Neuringer, Kornell et Olufs (2001) montrent que même appliquée après un programme de renforcement opérant qui a préalablement accru la variabilité, l'extinction entraîne une augmentation supplémentaire de cette variabilité.

Ensuite, davantage que le type de programmes utilisés (à ratio versus à intervalle par exemple), c'est la fréquence du renforcement appliqué dans ces programmes qui a une influence (non opérante) sur la variabilité des réponses (Lee et al., 2007; Machado, 1989; Neuringer, 2002, 2009; Page & Neuringer, 1985).

Cette conclusion est affirmée par des études qui examinent l'influence de la probabilité de renforcement sur la variabilité non plus au niveau global (celui de la session expérimentale ou de l'expérience), mais à un niveau plus local : celui de l'essai. Elles tirent profit du fait qu'au fil des essais, ou au fil du temps, la probabilité d'obtenir un agent renforçateur change.

Ainsi, Cherot, Jones et Neuringer (1996) ont inspecté chez des rats et des pigeons les effets d'un programme de renforcement à ratio fixe (où une séquence correcte sur quatre est renforcée) sur les séquences d'appuis sur des leviers ou de coups de bec sur des clés ; le renforcement est contingent soit à la variabilité, soit à la répétition d'une séquence particulière. Dans ce programme la probabilité de recevoir l'agent renforçateur est maximale lors de la dernière séquence. Cherot et al. (1996) observent que la variabilité comportementale diminue avec l'approche de l'agent renforçateur, que celui-ci soit contingent à la variabilité ou à la répétition des séquences.

Quant à Roberts et son équipe, ils testent différentes associations entre des stimulus colorés et des probabilités de renforcement contingent à la simple émission d'une réponse (quel que soit son niveau de variabilité). Ils constatent que plus cette probabilité est faible, plus la durée des appuis sur un levier chez des rats (Gharib, Gade, & Roberts, 2004) et plus la localisation des coups de bec chez des pigeons (Stahlman, Roberts, & Blaisdell, 2010) sont variables. La probabilité d'une récompense modifie la variabilité de dimensions qui ne font pas directement l'objet d'un programme de renforcement opérant.

Enfin, cette variabilité comportementale, répercussion indirecte des différents programmes de renforcement mis en place, peut interagir avec les contingences de renforcement opérant, direct, de la variabilité (Grunow & Neuringer, 2002; Lee et al., 2007; Machado, 1989, 1992; Neuringer, 2002, 2009).

C'est ce que montrent Grunow et Neuringer (2002) en manipulant indépendamment ces deux sources de variabilité. Ils constituent plusieurs groupes de rats chez qui ils induisent différents niveaux de variabilité, corrélés à la sévérité de critères de renforcement contingent à cette propriété. D'abord, ils appliquent des programmes de renforcement continu, puis ils réduisent progressivement le taux de renforcement. Grunow et Neuringer (2002) observent que l'intermittence du renforcement a des effets sur la variabilité différents selon les groupes. Elle entraîne une diminution de cette propriété lorsque le niveau de variabilité initial est élevé mais une augmentation lorsque le niveau de variabilité initial est faible (voir Maes, 2003, pour des résultats similaires obtenus chez l'être humain). Ces effets de la fréquence du renforcement sont cependant moins importants que ceux produits par le renforcement direct de la variabilité.

En premier lieu, ces résultats confirment donc que le renforcement opérant, direct, de la variabilité semble être une méthode à privilégier pour augmenter la variabilité comportementale. En effet, le niveau de variabilité peut alors être contrôlé. En outre le taux de réponses ne devrait pas subir de diminution, un effet observé lors de l'extinction (Antonitis, 1951; Neuringer et al., 2001) et de la diminution de la fréquence de renforcement (Grunow & Neuringer, 2002).

En second lieu, ces recherches soulignent la nécessité de contrôler la probabilité de renforcement. Elles mettent en évidence les limites des trois procédures de renforcement opérant de la variabilité présentées précédemment (*lag*, *least frequent* et *threshold procedures*). En effet, dans ces procédures, si l'apprentissage va dans le sens attendu, la probabilité d'émettre des réponses susceptibles d'être renforcées (c'est-à-dire de plus en plus variables) augmente. De plus, le taux de renforcement est souvent inversement corrélé à la sévérité des critères de renforcement.

Lorsqu'il n'est pas possible de maîtriser précisément le taux de renforcement, l'utilisation de groupes contrôles dans les expériences sur la variabilité opérante s'avère cruciale. Il s'agit d'égaliser les fréquences de renforcement entre le groupe expérimental et le groupe contrôle, mais tandis que le renforcement est contingent à la variabilité pour le premier, il ne l'est pas pour le second. Le groupe contrôle *yoke* répond précisément à cette exigence. Les moments où les sujets du groupe expérimental obtiennent des agents renforçateurs sont enregistrés et cet enregistrement est ensuite utilisé pour dispenser les mêmes stimulus aux sujets du groupe contrôle, indépendamment de la variabilité de leurs réponses. Bien sûr les niveaux de variabilité devraient être différents selon les groupes – davantage prédits et maîtrisés lorsque le renforcement est contingent à la variabilité. C'est le résultat qui a en effet été obtenu dans les différentes expériences sur la variabilité comportementale (ex. : Neuringer, 2002, 2009; Page &

Neuringer, 1985), et qui constitue un premier élément montrant que cette dimension dépend bien des contingences de renforcement.

2.3.3. La nature opérante de la variabilité comportementale

Bien évidemment, les recherches sur la variabilité opérante ne se bornent pas à signaler que les individus peuvent apprendre à diversifier leurs comportements. Elles attestent également la nature opérante de cette dimension en montrant qu'elle possède les mêmes caractéristiques que les autres dimensions opérantes.

L'une de ces caractéristiques porte sur le contrôle discriminatif (Denney & Neuringer, 1998; Page & Neuringer, 1985). Notamment, Denney et Neuringer (1998) comparent la variabilité obtenue lors de deux expériences, l'une consistant à appliquer un programme de renforcement multiple (dans lequel les composantes successives sont signalées par des stimulus discriminatifs), et l'autre un programme de renforcement mixte (qui se différencie du précédent dans la mesure où les stimulus discriminatifs sont enlevés). La première composante de chacun de ces programmes consiste à présenter des agents renforçateurs de manière contingente à la variabilité des réponses. Dans la seconde composante, le taux de renforcement est identique (grâce à une procédure *yoke*), mais le renforcement ne dépend pas de cette variabilité. Denney et Neuringer (1998) observent que dans le programme de renforcement multiple, le niveau de variabilité change immédiatement, dès la présentation des stimulus discriminatifs, et avant l'arrivée des agents renforçateurs, contrairement à l'autre programme. Cela confirme que la variabilité comportementale peut être contrôlée par des contingences à trois termes.

Une seconde caractéristique, relative au choix de produire des réponses variables versus stéréotypées, a été examinée (Abreu-Rodrigues, Lattal, dos Santos, & Matos, 2005; Neuringer, 1992). Néanmoins, les conclusions de ces études se heurtent à la différence entre les taux de renforcement obtenus après les différents types de séquences – le taux est plus élevé après les séquences répétitives. Dès lors, les auteurs concluent de manière générale et avec une relative prudence que les contingences de renforcement qui portent sur les niveaux de variabilité des réponses affectent le choix opéré par les sujets, comme c'est le cas pour d'autres comportements opérants. En effet, cela rappelle la loi du *matching* évoquée précédemment.

Un autre argument en faveur de la nature opérante de la variabilité comportementale est apportée par Machado (1989, 1992), qui montre que les valeurs de cette dimension peuvent être

situées sur un continuum, et ce en fonction des contingences de renforcement en vigueur – comme pour d'autres dimensions opérantes. Dans ses expériences, il apprend à des pigeons à émettre des séquences de coups de bec sur des clés ; les agents renforçateurs sont dispensés selon différents critères de variabilité. Dans son étude de 1989, il utilise quatre groupes de pigeons pour tester les effets de quatre critères ; dans celle de 1992, il teste les effets de deux critères différents en intra-sujets. Ses recherches sont d'autant plus convaincantes que sa procédure lui permet de manipuler les exigences des programmes de renforcement tout en égalisant le taux de renforcement à travers les groupes de sujets ou les conditions expérimentales.

Pour cela, il superpose à la procédure de renforcement des réponses les moins fréquentes (*least frequent procedure*) une procédure de *renforcement percentile* (Galbicka, 1994). Cette dernière consiste à dispenser les agents renforçateurs en fonction d'un critère de renforcement ordinal, qui reste constant pendant toute la session expérimentale. Par conséquent, la valeur cardinale de ce critère, elle, évolue à chaque essai, en fonction des performances du sujet. Plus précisément, Machado (1989, 1992) attribue à chaque séquence un score de récurrence, correspondant au nombre d'essais qui se sont écoulés depuis la dernière émission de cette séquence. Il peut ainsi obtenir une distribution des fréquences relatives de ces scores. C'est en manipulant le percentile en dessous duquel les séquences seront renforcées que Machado (1989, 1992) peut modifier les exigences de la variabilité requise par le programme de renforcement, tout en maintenant constante la probabilité de renforcement tout au long de l'expérience.

Il obtient effectivement des niveaux de variabilité de séquences corrélés aux exigences des programmes de renforcement, appuyant les corrélations observées par ailleurs, mais sans un tel contrôle du taux de renforcement (Abreu-Rodrigues et al., 2005; Grunow & Neuringer, 2002; Page & Neuringer, 1985).

Un certain nombre d'arguments, dont les principaux ont été présentés ci-dessus, a donc été accumulé en faveur de l'hypothèse selon laquelle la variabilité comportementale est fonctionnelle – du moins en partie. C'est à dire qu'une part de la variabilité comportementale dépend de ses propres conséquences : un niveau particulier de variabilité semble être lui-même *sélectionné*. Gharib et al. (2004) expriment cette idée en termes de coûts et bénéfices :

une trop grande déviation par rapport à ce qui a fonctionné peut faire échouer une action – la variabilité a alors un coût. D'un autre côté, essayer de nouvelles façons de faire les choses peut permettre de découvrir de meilleures façons de les faire – la variabilité a alors un bénéfice (p. 280, traduction par nos soins).

Selon eux, il existerait alors un « niveau optimal de variabilité », en partie déterminé par l'environnement – qui, lui, n'est ni fixe, ni incontrôlable.

Appliquée au domaine sensori-moteur, cette approche opérante de la variabilité diffère de celle des théories computationnelles du contrôle moteur postulant la nécessité de minimiser ce qu'elles considèrent uniquement comme le reflet d'un bruit interne aléatoire.

2.4. Hypothèses des recherches : le renforcement opérant, un mécanisme général d'apprentissage permettant de modifier la distribution des amplitudes des saccades oculaires

L'interaction entre la variabilité et la sélection comportementales est la base de la procédure opérante de façonnage (« *shaping* »). Cette procédure est largement utilisée dans les situations éducatives pour faire apparaître des comportements qui ont une probabilité d'occurrence très faible ou pour modifier la forme de certains comportements. Le façonnage est aussi souvent employé en laboratoire au début des expériences avec des animaux afin de leur apprendre à répondre aux clés ou aux leviers par exemple.

Le façonnage correspond au renforcement différentiel des approximations successives du comportement cible, jusqu'à ce que celui-ci soit émis de manière fréquente et stable (Skinner, 1953). Au début, il s'agit de renforcer une approximation lointaine du comportement cible. Lorsque le critère de réussite est atteint pour cette première approximation, les critères de renforcement changent de manière à renforcer une approximation plus proche du comportement terminal. Quant à la première approximation, elle est placée sous extinction. C'est grâce à la variabilité comportementale que des approximations différentes peuvent apparaître, et que la sélection de l'approximation suivante peut avoir lieu.

Bien que l'emploi du façonnage soit largement répandu, il n'a fait l'objet d'aucune étude systématique à ce jour. Il existe un consensus sur quelques règles, présenté succinctement dans les manuels destinés principalement à la formation des psychologues comportementalistes (ex. : Cooper, Heron, & Heward, 2007; Domjan, 2009), et rappelées par Galbicka (1994). Ainsi, après avoir spécifié le comportement cible, il s'agit de mesurer le niveau initial de performance. Puis il convient de déterminer les étapes intermédiaires du façonnage : elles doivent changer graduellement, afin d'une part d'éviter d'éteindre toutes les réponses (s'il s'avère difficile de passer d'une étape à une autre) ou de rendre la progression très lente (si les changements de critères sont trop minimes). De même, la durée des étapes doit être suffisante pour qu'une approximation soit

acquise, mais pas excessive non plus pour éviter que l'individu n'émette plus que cette approximation.

Il reste que l'application du façonnage varie en fonction des individus (ceux qui mettent en place la procédure et ceux qui y sont soumis), des tâches à apprendre, des environnements, si bien que certains auteurs comme Galbicka (1994) ou Lattal et Neef (1996) ont employé le terme « d'art » – la procédure de renforcement percentile faisant figure d'exception. C'est cette procédure de renforcement percentile que nous avons employée pour modifier des distributions d'amplitudes saccadiques.

La première étude teste l'hypothèse selon laquelle la modification du gain des saccades oculaires relève d'un processus général d'apprentissage – un apprentissage opérant. Si tel est le cas, l'utilisation d'une procédure de renforcement telle que le façonnage devrait permettre d'induire des changements de gain comparables à ceux produits par la procédure conventionnelle de double saut. Selon la théorie de l'apprentissage opérant, la modification des comportements s'explique par un changement dans les contingences, c'est-à-dire un changement dans les relations entre les comportements et leurs conséquences. Ces dernières ne sont pas définies a priori par leur forme mais avant tout par leur *fonction* – elles déterminent le comportement. Le changement de gain saccadique devrait donc s'observer quelle que soit la nature des conséquences des saccades, ce qui est testé dans plusieurs expériences par l'utilisation de différentes conséquences, dont la forme est sonore et/ou visuelle.

Si le contrôle et l'apprentissage oculo-moteurs dépendent effectivement d'un mécanisme général d'apprentissage, cela signifie qu'ils n'impliquent pas nécessairement des mécanismes spécialisés qui, comme le postule l'approche computationnelle classique, seraient responsables par exemple de la calibration d'un gain interne ou encore de la minimisation de la variabilité. Cette dernière doit elle aussi pouvoir s'expliquer (en partie du moins) par l'influence de ses propres conséquences fonctionnelles. L'objectif des deux recherches suivantes est de montrer qu'une part de la variabilité de l'amplitude saccadique résulte d'un apprentissage opérant.

Dans la deuxième étude, il s'agit d'induire expérimentalement différents niveaux de variabilité du gain de saccades horizontales en manipulant les contingences de renforcement : les saccades dont l'amplitude répond aux critères de variabilité requise sont suivies par un son.

L'utilisation d'agents renforçateurs auditifs met l'accent sur la fonction des conséquences des mouvements oculaires dans le contrôle de la variabilité des saccades. La nature de cette conséquence a été choisie sur la base des résultats des expériences antérieures. Or la valeur d'une

conséquence se détermine à l'aune de ses effets sur le comportement. La conclusion précédente n'équivaut donc pas à dire qu'une conséquence visuelle peut effectivement contrôler la variabilité de l'amplitude des saccades. Cette hypothèse est testée dans une troisième étude : un nouveau paradigme expérimental impliquant une tâche de recherche visuelle permet d'examiner si le fait de trouver une cible parmi des distracteurs permet de renforcer différents niveaux de variabilité d'amplitudes saccadiques.

ETUDE 1 : MODIFICATION DU GAIN
SACCADIQUE PAR LES CONTINGENCES
DE RENFORCEMENT

Afin de tester l'hypothèse selon laquelle la modification du gain des saccades oculaires relève d'un processus général d'apprentissage, plusieurs expériences ont été réalisées dans cette étude, publiée en 2011 dans le *Journal of Neurophysiology* (l'article correspondant est inséré dans la dernière annexe de ce manuscrit). Nous y avons manipulé les conséquences des saccades afin de mettre en place des procédures opérantes de façonnage.

Dans les expériences 1 et 2, la cible était éteinte au début de la saccade si bien que l'erreur rétinienne post-saccadique était absente. Un stimulus sonore était émis immédiatement après l'exécution des saccades dont l'amplitude correspondait aux critères de renforcement. Dans la mesure où les procédures de renforcement se sont avérées efficaces pour induire une diminution (expérience 1) et une augmentation (expérience 2) de l'amplitude des saccades, nous avons investigué d'autres caractéristiques de cette adaptation saccadique induite par renforcement opérant.

Ainsi, dans l'expérience 3, les saccades satisfaisant aux critères de renforcement étaient suivies de l'apparition de la cible à l'endroit où l'œil atterrissait. En revanche, les saccades dont l'amplitude ne correspondait pas à ces critères n'étaient suivies d'aucune conséquence.

Enfin, dans les expériences 4 et 5, une autre forme de conséquences a été testée. Il s'agissait de faire réapparaître la cible en vision centrale après chaque saccade, donc d'annuler l'erreur visuelle post-saccadique mais sans supprimer la stimulation visuelle – puisque celle-ci dépend de la distance entre la cible et la fovéa à la fin de la saccade. Un stimulus sonore était à nouveau utilisé pour renforcer les saccades dont l'amplitude répondait aux exigences de la procédure d'apprentissage. Dans cette expérience, nous avons également comparé certaines spécificités des adaptations saccadiques obtenues par le paradigme conventionnel de double saut (McLaughlin, 1967) et par les procédures de renforcement, notamment en examinant le transfert de ces adaptations à des saccades d'amplitudes et de directions différentes.

I. Méthode

1.1. Participants

Sur les 21 participants, 19 étaient naïfs au regard de l'objectif de cette étude. Ils n'avaient réalisé aucune expérience sur les mouvements oculaires auparavant, contrairement à deux des trois auteurs (S1 et S2), qui ont passé les expériences 1, 3 et 4. Comme il sera détaillé ci-dessous, certains participants ont aussi effectué plusieurs des expériences ; dans ce cas elles étaient espacées d'au moins sept jours. Les participants étaient âgés de 18 à 35 ans, ils avaient tous une vision normale ou corrigée. Ils ont signé un formulaire de consentement, dont un exemplaire est inséré en annexe 1.

Afin de se familiariser avec le matériel et la procédure de calibration, ils ont réalisé une session de 200 saccades horizontales vers une cible apparaissant 14.2° à droite d'un point de fixation. Ces données ne sont pas retranscrites.

1.2. Equipement et stimulus

Les participants étaient assis face à un moniteur (Iiyama HM204DT, 100 Hz) qui affichait les différents stimulus visuels. Leur tête était stabilisée par un dispositif de maintien de la mâchoire supérieure et une barre frontale de telle sorte que leurs yeux étaient situés à 60 cm du centre de cet écran.

Les stimulus étaient générés par un Power Mac G4 à l'aide de l'extension de la *Psychophysics Toolbox* pour Matlab. La présentation des stimulus, ainsi que l'enregistrement, l'affichage et le stockage des réponses étaient contrôlés par un boîtier PXI, via une interface réalisée grâce au logiciel LabView Real-Time (National Instrument). La synchronisation entre l'enregistrement des données et l'affichage des stimulus se faisait à une résolution de 1 ms grâce à des signaux provenant de l'ordinateur affichant les stimulus. Les mouvements oculaires étaient mesurés continuellement par une caméra à infrarouges (ISCAN, RK-276) à 240 Hz. Avant chaque session, les participants devaient effectuer une calibration de l'appareil en fixant tour à tour, plusieurs fois, sept points distribués sur la ligne horizontale traversant le centre de l'écran. Cela permettait

d'obtenir une fonction linéaire servant à convertir les valeurs brutes issues du système d'enregistrement en positions de l'œil.

Une croix de fixation (un signe ' + ', blanc, de 0.5° de hauteur et de largeur) marquait le début d'un essai. La cible avait la forme d'un patch gaussien (écart-type : 0.93°, luminance maximum : 62 cd/m²). Ces deux stimulus apparaissaient sur un fond gris uniforme (luminance : 40 cd/m²). Quant au son utilisé comme agent renforçateur à la fin des saccades, il correspondait à une onde sinusoïdale de 1000 Hz et d'une durée de 60 ms.

1.3. Le façonnage par procédure de renforcement percentile ciblé

Afin d'induire des modifications de l'amplitude des saccades, il était nécessaire de contrôler les conséquences de ces mouvements. Pour cela, lorsque les participants effectuaient une saccade vers la cible (apparaissant 14.2° à droite de la croix de fixation), cette dernière était éteinte au début du mouvement de l'œil. Les saccades dont l'amplitude correspondait à des critères de renforcement spécifiques, déterminés par rapport aux essais précédents, étaient suivies d'une conséquence sonore (expériences 1, 2, 4 et 5) et/ou visuelle (expériences 3, 4 et 5) en fonction de l'expérience considérée – rappelons que pour les deux dernières expériences, la cible apparaissait en vision centrale après chaque saccade.

Les critères de renforcement étaient déterminés selon une procédure de renforcement percentile ciblé (appelée « *targeted percentile schedule* », utilisée par Galbicka, 1994; Galbicka, Kautz, et Jagers, 1993, et Madelain et al., 2007). Cette procédure de renforcement différentiel (c'est-à-dire consistant à renforcer certains comportements tandis que d'autres étaient éteints) avait pour objectif de centrer la médiane des amplitudes autour d'une valeur précise (18.9° dans l'expérience 2, 9.4° dans les autres expériences, c'est-à-dire une différence de plus ou moins 33% par rapport l'amplitude du saut initial de la cible).

Le programme de renforcement percentile (Galbicka, 1994, voir aussi Machado, 1989, 1992) repose sur la possibilité d'affecter un rang à chaque nouvelle réponse, en la situant parmi un certain nombre n de réponses précédentes ordonnées. Pour qu'une nouvelle réponse soit considérée comme correcte, elle doit non pas dépasser une valeur fixée à l'avance par l'expérimentateur, mais être meilleure que la valeur prise par la réponse antérieure (parmi les n réponses précédentes) située à un rang k déterminé (le percentile). Autrement dit, la valeur

cardinale du critère de renforcement n'est pas établie a priori, elle correspond à la valeur prise par la $k^{ème}$ réponse.

Le renforcement percentile ciblé consiste à mettre en œuvre deux programmes de renforcement percentile simultanément, l'un s'appliquant aux réponses situées en dessous de la valeur cible, l'autre s'appliquant à celles qui sont situées au dessus. Comme son nom l'indique, ce programme permet a priori de stabiliser les réponses autour de cette valeur cible.

Dans nos expériences, la probabilité de renforcement était fixée à 40% des saccades (20% pour les valeurs situées au dessus de la valeur cible, autant pour les valeurs en dessous). La figure 7 A illustre la méthode utilisée pour déterminer l'intervalle des amplitudes renforcées pour une saccade, lorsque la valeur cible était de 9.4°.

L'amplitude de la nouvelle saccade était comparée aux valeurs des 50 saccades précédentes, une fois ordonnées. D'une part, pour que la saccade soit renforcée, son amplitude devait être supérieure à celle de 80% des saccades courtes (c'est-à-dire des saccades comprises entre 0 et l'amplitude cible). Sur la figure 7 A, elle devait être supérieure à 8.8°. D'autre part, l'amplitude de la saccade devait être inférieure à celle de 80% des saccades longues (celles dont l'amplitude était supérieure à la valeur requise). Dans l'exemple, cette limite était de 9.8°.

La courbe représentée en gris clair sur cette figure représente la distribution des amplitudes cumulées obtenues à partir de l'une des premières sessions de façonnage, lorsque la médiane de la distribution était nettement supérieure à la valeur cible. Par conséquent il n'y avait pas suffisamment de valeurs inférieures à 9.4°. Dans ce cas toutes les saccades dont l'amplitude était inférieure à 9.4° étaient renforcées. De manière similaire, s'il n'y avait pas suffisamment de valeurs supérieures à la valeur cible, toutes les saccades dont l'amplitude était supérieure à la valeur cible étaient renforcées – seules ces situations donnaient lieu à une probabilité de renforcement supérieure à 40%.

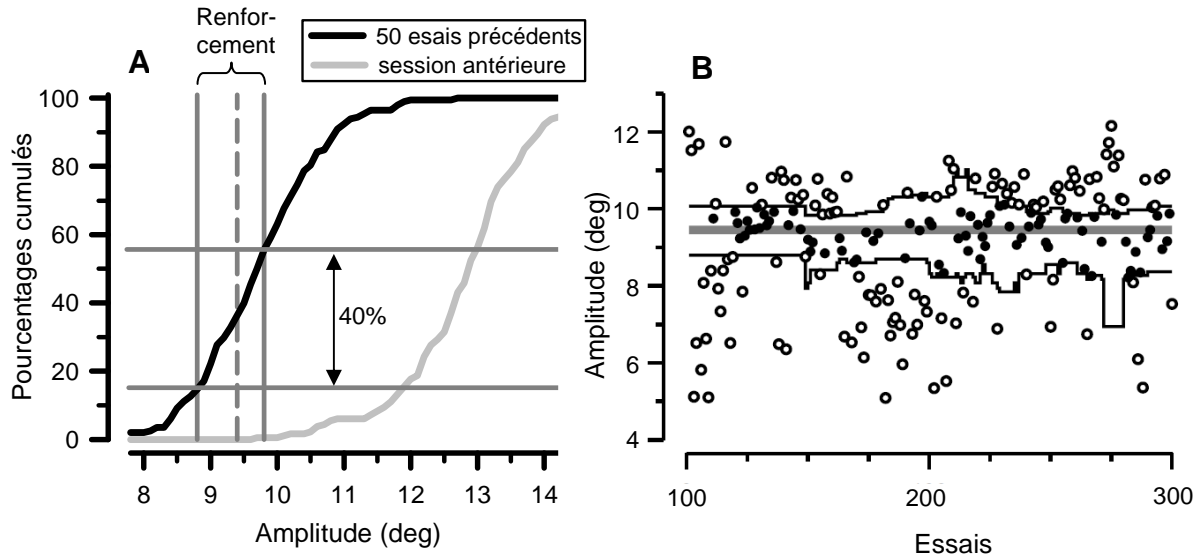


Figure 7. Procédure de renforcement percentile ciblé lors d'un façonnage visant à réduire le gain saccadique. A : sur la distribution des amplitudes cumulées des 50 essais précédents (courbe noire), les 40% des saccades les plus proches de la valeur cible (9.4° , trait vertical en pointillés) déterminent l'intervalle d'amplitudes renforcées (traits pleins verticaux) pour la saccade qui va être exécutée. En gris : distribution des fréquences cumulées obtenues pour une session antérieure. B : Changement des critères de renforcement (traits noirs) au fil des essais d'une session d'apprentissage pour le participant S2 (expérience 3). La ligne grise horizontale correspond à l'amplitude cible, les cercles noirs aux saccades renforcées et les cercles blancs aux saccades non renforcées.

Lors de la première session de façonnage, les limites de l'intervalle de renforcement étaient calculées à partir des saccades de la ligne de base préalable. Lors des premiers essais de chaque session, elles étaient calculées à partir des sessions précédentes. En effet, les changements d'amplitudes observés se poursuivaient d'une session à l'autre, ce qui permettait de renforcer une proportion déterminée de saccades dès le début de la session, comme le reflète la figure 7 B. Dans cet exemple, la cible effectuait un saut initial de 14.2° et la valeur visée par le renforcement était de 9.4° ; la figure illustre les ajustements essai après essai des critères de renforcement, ceux-ci dépendant à la fois de la valeur cible et des amplitudes des 50 saccades précédentes.

Cette procédure présente plusieurs avantages. Tout d'abord, comme nous l'avons mentionné antérieurement, la probabilité de renforcement est fixée par l'expérimentateur et reste globalement constante pendant toute la session. Ensuite, à chaque essai, une saccade peut être

potentiellement renforcée. Enfin, une fois que le comportement est façonné, autrement dit que l'amplitude des saccades s'approche de la valeur cible, cette procédure amène le comportement de plus en plus près de ce qui est attendu, réduisant les écarts entre les réponses au fil des essais.

Toutefois, parce que justement ce programme est construit de telle sorte qu'il « permette [au critère de renforcement] de revenir vers un moins bon score si le comportement se déplace systématiquement dans cette direction » (Galbicka, 1994, p. 749 ; voir aussi Galbicka et al., 1993) sans que cela n'affecte la probabilité de renforcement, des réponses relativement éloignées du comportement cible peuvent être suivies d'un agent renforçateur. Par conséquent, les participants pourraient adopter une stratégie globale consistant à maintenir constante la distribution de leurs réponses, sans la déplacer vers la valeur cible – ce qui n'induirait pas de diminution de taux de renforcement.

Cela peut être évité si l'individu maximise le taux local de renforcement, ce qui a davantage de chances de se produire si les agents renforçateurs ont une valeur renforçante élevée. Pour cela, dès que trois essais sur cinq étaient renforcés, les participants entendaient une voix synthétique prononcer « bravo ! » ; indiquant qu'ils avaient gagné un point – le fait que ce procédé ait permis de contrôler le comportement des participants de l'expérience de Madelain et al. (2007) laissait présager de la valeur renforçante de ces conséquences ; en outre cela a été confirmé a posteriori dans cette étude, tout comme dans la suivante.

Les participants avaient pour consigne de « regarder la croix, puis la cible et de gagner le plus de points possible », comme dans un jeu (l'intégralité des consignes figure en annexe 2). Ils étaient informés qu'ils étaient en compétition avec les autres participants dont ils pouvaient connaître les scores, le vainqueur recevant un lot à la fin de l'expérience.

1.4. Procédures des différentes expériences

1.4.1. Expérience 1 : façonnage d'une diminution du gain saccadique avec agent renforçateur auditif

Cette expérience consistait à tester l'hypothèse selon laquelle il est possible de renforcer la diminution de l'amplitude des saccades en l'absence de signal d'erreur rétinienne post-saccadique. Les saccades dont l'amplitude était comprise dans l'intervalle de renforcement étaient renforcées par un son.

Quatre personnes (S1, S2, S5 et S6) ont participé à cette expérience, qui comprenait quatre conditions successives : une *ligne de base conventionnelle*, une *ligne de base renforcée*, la condition d'*adaptation* et enfin les sessions de *désadaptation*. Notons que dans un souci de concision, les termes « d'adaptation » et « désadaptation » seront conservés dans le cadre des différents façonnages pour faire référence à l'application de différents programmes de renforcement percentiles ciblés, même si ces procédures ne correspondent pas à celles employées dans les paradigmes conventionnels de double saut.

a. Ligne de base conventionnelle

Lors de cette première condition expérimentale, les participants avaient pour tâche de réaliser 200 saccades vers une cible stationnaire. Aucun son n'était émis après les saccades.

Chaque essai commençait par l'apparition de la croix de fixation à 7.1° à gauche du centre de l'écran. Puis, entre 750 et 1000 ms plus tard, ce point disparaissait et simultanément, la cible apparaissait à droite, à une excentricité de 14.2° ; elle restait à cette position pendant 1000 ms. Seule la première saccade émise lors de ces 1000 ms était enregistrée.

b. Ligne de base renforcée

Cette seconde ligne de base consistait à émettre 200 saccades, initiées après la disparition de la croix de fixation et dirigées vers la cible qui, comme dans la condition précédente, apparaissait 14.2° à droite de ce point. Mais dans cette condition expérimentale, la cible disparaissait au début de la saccade et la procédure de renforcement percentile ciblé était appliquée – l'amplitude visée était 14.2° . Les saccades dont l'amplitude correspondait aux critères de renforcement étaient suivies d'un son indiquant que l'essai était correct. Rappelons que, pour les 50 premiers essais, les limites de l'intervalle des amplitudes renforcées étaient calculées à partir de la session précédente, à savoir la ligne de base conventionnelle.

c. Adaptation

Ces sessions étaient identiques à celles de la condition précédente, excepté que le programme de renforcement percentile ciblait une valeur de 9.4° (une diminution de 33% de l'excentricité de

la cible apparaissant en vision périphérique). Le nombre de sessions d'adaptation effectuées dépendait des participants, et ce quelle que soit l'expérience. En général, ils réalisaient deux sessions par jours, séparées par une pause de 10 à 20 minutes pendant lesquelles ils étaient libres de leurs mouvements.

d. Désadaptation

Après les sessions d'adaptation, les participants ont accompli une à trois sessions de désadaptation similaires aux essais de la ligne de base renforcée, c'est-à-dire avec une amplitude cible de 14.2° , un agent renforçateur sonore et en l'absence d'erreur rétinienne post-saccadique.

1.4.2. Expérience 2 : façonnage d'une augmentation du gain saccadique avec agent renforçateur auditif

Les conditions expérimentales, passées par quatre autres participants (S9, S10, S11 et S21), étaient identiques à celles de l'expérience précédente, excepté la valeur cible de la condition d'adaptation. En effet, le programme de renforcement percentile ciblait cette fois une amplitude de 18.9° (une augmentation de 33% de l'excentricité de la cible).

Seuls deux des participants (S9 et S21) ont accompli les sessions de désadaptation (où la valeur cible était ramenée à 14.2°).

1.4.3. Expérience 3 : façonnage d'une diminution du gain saccadique avec agent renforçateur visuel (la réapparition de la cible en vision centrale)

Cette expérience avait pour objectif de tester la possibilité de réduire le gain saccadique via une procédure de façonnage dans lequel l'agent renforçateur était visuel – cette fois aucun son n'était délivré à l'issue des saccades. Comme dans la première expérience, la valeur cible des sessions d'adaptation était de 9.4° .

L'agent renforçateur correspondait à l'apparition de la cible en vision centrale, c'est-à-dire que la cible réapparaissait de telle sorte qu'elle soit alignée avec la fovéa (le terme de *stabilisation de l'image de la cible* est employé, le procédé mis en œuvre est détaillé plus bas). Il n'y avait donc pas d'erreur de position rétinienne. La cible restait visible jusqu'à la fin de l'essai. En revanche, à

l'issue des saccades dont l'amplitude ne correspondait pas à l'intervalle requis, la cible ne réapparaissait pas.

Quatre individus (S1, S2, S3, et S4) ont passé cette expérience, y compris la condition de désadaptation.

1.4.4. Expérience 4 : façonnage d'une diminution du gain saccadique avec agent renforçateur auditif et stabilisation de l'image de la cible à l'issue de chaque essai

Le façonnage avec agent renforçateur auditif serait-il efficace malgré la présence de l'image de la cible sur la fovéa à l'issue de chaque essai ? Pour répondre à cette question, cette expérience a consisté à utiliser des agents renforçateurs sonores après les saccades dont l'amplitude correspondait aux critères du façonnage, comme dans l'expérience 1, mais à faire réapparaître la cible en vision centrale *quelle que soit l'amplitude des saccades*. Autrement dit, l'erreur de position rétinienne était systématiquement annulée et seule l'émission du son était contingente à l'amplitude saccadique.

Ce sont les participants S1, S2, S7 et S8 qui ont effectué les quatre conditions expérimentales.

1.4.5. Expérience 5 : comparaison entre les adaptations saccadiques induites par la procédure de renforcement et par le paradigme conventionnel

Afin de comparer l'adaptation saccadique induite par le paradigme conventionnel de double saut et celle obtenue par façonnage opérant, il s'agissait de faire passer ces deux paradigmes aux mêmes participants et de comparer certaines des caractéristiques de l'adaptation saccadique : le transfert à d'autres amplitudes et à d'autres positions, ainsi que la désadaptation.

a. Adaptations du gain saccadique

Que ce soit lors de l'adaptation induite par le façonnage ou par la procédure de double saut, les participants effectuaient d'abord une ligne de base conventionnelle : 200 saccades émises entre la croix de fixation située à 7.1° à gauche du centre de l'écran et la cible apparaissant 14.2° vers à droite – elle restait visible jusqu'à la fin de l'essai.

Puis, dans la procédure de façonnage, une session de ligne de base renforcée avait lieu (200 essais, avec une valeur cible de 14.2°), suivie de 6 sessions de 200 essais d'adaptation (la valeur cible étant alors de 9.4°). Les agents renforçateurs utilisés étaient les sons, et l'image de la cible était systématiquement stabilisée sur la fovéa après chaque essai, comme dans l'expérience 4.

Quant au paradigme de double saut, les participants ont réalisé 2 sessions de 200 essais pendant lesquels la cible effectuait un saut intra-saccadique de 4.8° dans le sens opposé à la direction de la première saccade (elle réapparaissait donc à une excentricité de 9.4° par rapport à la position occupée par la croix de fixation), dans le but d'induire une diminution du gain.

b. Transfert

Après avoir achevé les sessions d'adaptation, les participants accomplissaient une session visant à tester le transfert d'adaptation. Ce type de sessions commençait par 100 essais d'adaptation (soit via le renforcement sonore, soit via le double saut de la cible, selon le paradigme précédemment appliqué), puis les participants devaient exécuter 150 essais pendant lesquels la cible était éteinte au début du mouvement. Dans le cas d'une adaptation par façonnage préalable, aucun son n'était émis à l'issue de ces 150 saccades.

Lors de ces derniers essais, soit la position initiale de la croix de fixation (dans les essais dits de « transfert de position »), soit l'excentricité de la cible (dans les essais de « transfert d'amplitude ») étaient manipulées. Dans le premier cas, la croix de fixation pouvait apparaître à trois positions différentes : à 4.7° , 7.1° ou 9.4° à gauche du centre de l'écran. La cible apparaissait 14.2° à droite de cette croix. Dans le second cas, celui du transfert d'amplitude, la croix de fixation initiale apparaissait systématiquement à 7.1° à gauche du centre de l'écran, mais l'excentricité de la cible (apparaissant à droite) variait et pouvait être de 11.8° , 14.2° ou 16.5° . Les cinq types d'essais (les essais dont la croix de fixation initiale apparaissait à 7.1° et l'excentricité de la cible était de 14.2° étant communs aux deux types de transferts) étaient présentés 30 fois chacun, dans un ordre pseudo-aléatoire.

c) Désadaptation conventionnelle

Enfin, les participants réalisaient une session de désadaptation qui consistait en 100 essais d'adaptation (soit via le renforcement sonore, soit via le double saut de la cible selon le paradigme appliqué) suivis de 150 essais pendant lesquels la cible apparaissait à une excentricité de 14.2° à

droite de la croix de fixation. Cette dernière apparaissait toujours 7.1° à gauche du centre de l'écran. Quel que soit le paradigme appliqué, la cible n'effectuait pas de second saut et restait visible jusqu'à la fin de l'essai.

Quatre personnes (S12, S13, S14 et S15) ont passé cette expérience avec les deux paradigmes, toutes ont commencé par la procédure de façonnage et terminé par celle de double saut. Les passations de ces deux procédures expérimentales étaient espacées d'une semaine.

En outre, trois autres participants (S3, S16 et S17) ont passé uniquement le paradigme de double saut de la cible visant à réduire le gain saccadique. Enfin, les participants S18, S19 et S20 ont passé seulement le paradigme conventionnel d'adaptation saccadique visant à augmenter l'amplitude des saccades. La cible effectuait alors un saut intra-saccadique l'amenant à 18.9° à droite de la croix de fixation.

1.5. Acquisition et analyse des données

Seules les saccades qui apparaissaient dans une fenêtre de 1000 ms après l'apparition de la cible étaient détectées, mesurées *online* et prises en compte ensuite dans les analyses de données *offline*.

Les critères de renforcement étaient déterminés à partir des mesures *online*. Un algorithme temps réel se basait sur le dépassement d'un certain seuil de vitesse pour identifier le début et la fin de ces saccades. Dans les expériences 3 à 5, la cible disparaissait au début du mouvement de l'œil et elle réapparaissait de manière à être alignée avec la fovéa à la fin de la saccade ; elle restait alors à cette position jusqu'à la fin de l'essai. Pour effectuer cette stabilisation de l'image de la cible sur la fovéa, la position médiane de l'œil était calculée à partir des positions de l'œil relevées sur 12 ms à la fin de la saccade ; la cible était affichée à la position correspondante dès que le calcul était achevé. Les comparaisons des mesures *online* et *offline* ont indiqué que la cible était placée en moyenne à 0.17° (écart-type : 0.31°) de la position de l'œil et qu'à partir du début de la saccade, la cible était invisible pendant 65 ms en moyenne.

Pour l'examen *offline* des enregistrements, un programme interactif a été élaboré afin de filtrer, afficher et analyser les données. La vitesse et l'accélération horizontales étaient obtenues par différenciation des signaux de positions de l'œil (sur une fenêtre de 8 ms). Un logiciel écrit sous LabView permettait de détecter les saccades à partir de critères fixes de vitesse et d'accélération (respectivement de $30^{\circ}/s$ et $7000^{\circ}/s^2$). L'expérimentateur devait accepter ou rejeter

les mouvements repérés, ou le cas échéant ajuster le moment de leur initiation. L'amplitude des saccades correspondait à la différence entre les médianes des positions de l'œil calculées sur des fenêtres de 25 ms avant et après les saccades. Les saccades qui étaient émises dans la direction opposée à celle de l'apparition de la cible ou qui avaient une latence inférieure à 100 ms étaient écartées.

Plusieurs calculs ont été réalisés afin d'étudier les évolutions des gains saccadiques observées au cours des différentes expériences de façonnage, à savoir :

- La moyenne des amplitudes – qu'il fallait toujours diviser par 14.2 pour obtenir la moyenne des gains des gains puisque l'excentricité de la cible (avant sa disparition et son second saut éventuel) était fixe – des 50 dernières saccades de chaque condition expérimentale,
- les courbes de régression non linéaire des amplitudes en fonction des essais, avec un lissage obtenu par la technique du Lowess (sur des fenêtres de 100 essais),
- la proportion de changement de gain obtenue à la fin de l'adaptation par façonnage : l'amplitude moyenne des 50 derniers essais de la ligne de base renforcée (où l'amplitude cible était de 14.2°) a été comparée à celle des 50 derniers essais de la dernière session de la condition d'adaptation (où l'amplitude cible était soit 9.4°, soit 18.9°). Le pourcentage de changement d'amplitude était alors obtenu grâce à la formule :

$$\frac{\text{amplitude moyenne de l'adaptation} - \text{amplitude moyenne de la ligne de base}}{\text{amplitude moyenne de la ligne de base}} \times 100, \quad (1)$$

ce pourcentage étant équivalent à la proportion de changement de gain.

- la proportion de changement de gain obtenue à la fin de la désadaptation par façonnage, par rapport aux saccades de la condition d'adaptation antérieure. Cette proportion était calculée sur les 50 derniers essais de ces deux conditions expérimentales grâce à la formule :

$$\frac{\text{amplitude moyenne de la désadaptation} - \text{amplitude moyenne de l'adaptation}}{\text{amplitude moyenne de l'adaptation}} \times 100, \quad (2)$$

- enfin, les propriétés saccadiques telles que le pic de vitesse et la latence ont été également examinées.

Afin d'évaluer les changements induits par les différentes procédures, des comparaisons par paires a priori ont été réalisées entre les moyennes calculées sur les 50 derniers essais de différentes conditions. Ces comparaisons ont été effectuées sujet par sujet, via des tests *t*

bilatéraux soumis à la correction de Dunn-Sidak – ces tests ont été réalisés afin de répondre à la requête d'un des relecteurs de l'article au moment de sa publication (notre analyse initiale via la méthode du Bootstrap ayant été refusée) ; ils sont repris ici dans un souci de cohérence avec les données publiées. Il en va de même pour les tests t appariés utilisés dans la discussion de cette étude pour comparer les propriétés des saccades, tests réalisés sur un petit nombre de participants.

II. Résultats

De manière générale, les procédures de renforcement, que ce soit via des conséquences auditives et/ou visuelles, ont induit des modifications du gain des saccades, et ce de manière progressive.

2.1. Expérience 1 : façonnage d'une diminution du gain saccadique avec agent renforçateur auditif

Dans cette expérience, la procédure de renforcement percentile appliquée lors des sessions d'adaptation ciblait une amplitude de 9.4° (soit un gain de 0.66). Les données individuelles du participant S6 sont représentées sur la figure 8 A. Lors de la ligne de base conventionnelle, le gain saccadique moyen était de 0.93, conforme aux données rapportées par la littérature. Cette valeur était de 0.94 à la fin de la ligne de base renforcée, dans laquelle le programme de renforcement percentile ciblait un gain de 1. Pendant les sessions d'adaptation renforcée, le gain a diminué de manière progressive, pour parvenir à une valeur moyenne de 0.70 sur les 50 derniers essais.

Pour ce participant, l'amplitude cible a été atteinte en quatre sessions d'adaptation. L'écoulement d'une journée entre les sessions d'adaptation (représenté par la ligne grise verticale au niveau de l'essai 800) n'a donné lieu qu'à une perte minimale de l'apprentissage.

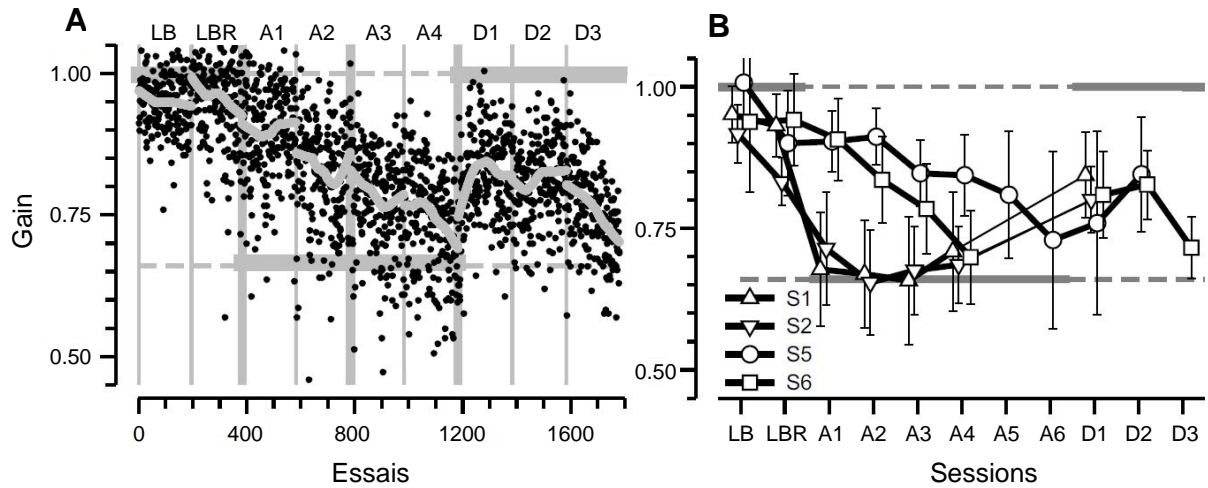


Figure 8. Evolution du gain saccadique au cours de l'expérience 1 (façonnage de la diminution du gain avec agent renforçateur auditif). A : gain saccadique observé au fil des essais pour l'individu S6. Chaque point représente le gain d'un seul essai, les courbes grises correspondent aux courbes de régression Lowess et les lignes grises verticales délimitent les sessions, les droites plus épaisses représentent l'écoulement d'une journée entre deux sessions (dont le type est indiqué par les lettres en haut du graphique, à savoir, LB : ligne de base conventionnelle, LBR : ligne de base renforcée, A : adaptation et D : désadaptation). B : gain moyen (et écart-type) calculé sur les 50 derniers essais de chaque session pour les quatre participants (les lignes fines entre les points relie la dernière session d'adaptation à la première session de désadaptation).

Comme le montre la figure 8 B, le gain des quatre participants de cette expérience a diminué au cours des sessions d'adaptation. Pour deux d'entre eux, il avait atteint le gain cible (0.66) à l'issue de la deuxième session, tandis que pour les deux autres, quatre sessions (pour S6) ou six sessions (pour S5) ont été nécessaires. La proportion de changement de gain lors de l'adaptation, par rapport à la ligne de base, était de -23.1% en moyenne. En général, lors de la désadaptation, le gain a augmenté par rapport à l'adaptation, excepté pour le participant S6 pour qui l'amplitude a diminué lors de la dernière session.

La figure 9 permet de comparer les changements de gain obtenus à la fin des différentes conditions expérimentales. Les comparaisons par paires (les crochets en haut de la figure) ont confirmé la réduction significative du gain observée lors des sessions d'adaptation pour tous les

participants ainsi que l'augmentation lors des sessions de désadaptation – sauf pour S6 dont le gain n'a pas augmenté lors de la désadaptation. Enfin, le gain a diminué significativement entre la ligne de base conventionnelle (pendant laquelle la cible était stationnaire) et la ligne de base renforcée (quand la cible disparaissait pendant la saccade) pour deux des quatre participants (S2 et S5, tests t corrigés, $p < 0.05$).

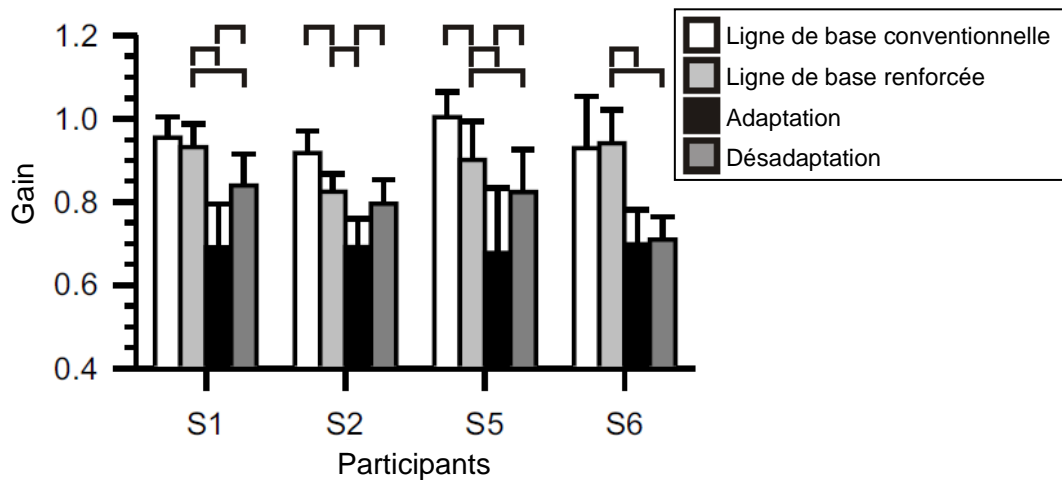


Figure 9. Comparaison des gains moyens (et écarts-types) calculés sur les 50 derniers essais de chaque condition expérimentale pour les quatre participants de l'expérience 1 (façonnage de la diminution du gain avec agent renforçateur auditif). Les crochets indiquent les différences significatives obtenues lors des tests t corrigés de comparaisons par paires a priori ($p < 0.05$).

2.2. Expérience 2 : façonnage d'une augmentation du gain saccadique avec agent renforçateur auditif

Cette fois, la procédure de renforcement percentile encourageait une augmentation du gain. La figure 10 A illustre l'évolution du gain au fil des essais pour le participant S9. Pendant les lignes de base conventionnelle et renforcée, le gain moyen était respectivement de 0.98 et 0.92. L'augmentation a été visible lors de la quatrième session d'adaptation ; à la fin de cette condition expérimentale le gain moyen était de 1.07. Pour ce participant, la proportion de changement de gain entre la ligne de base et l'adaptation était donc de 19.7%. Entre l'adaptation et la

désadaptation, cette valeur était de -12.9% : le gain a donc nettement diminué suite au changement des contingences de renforcement.

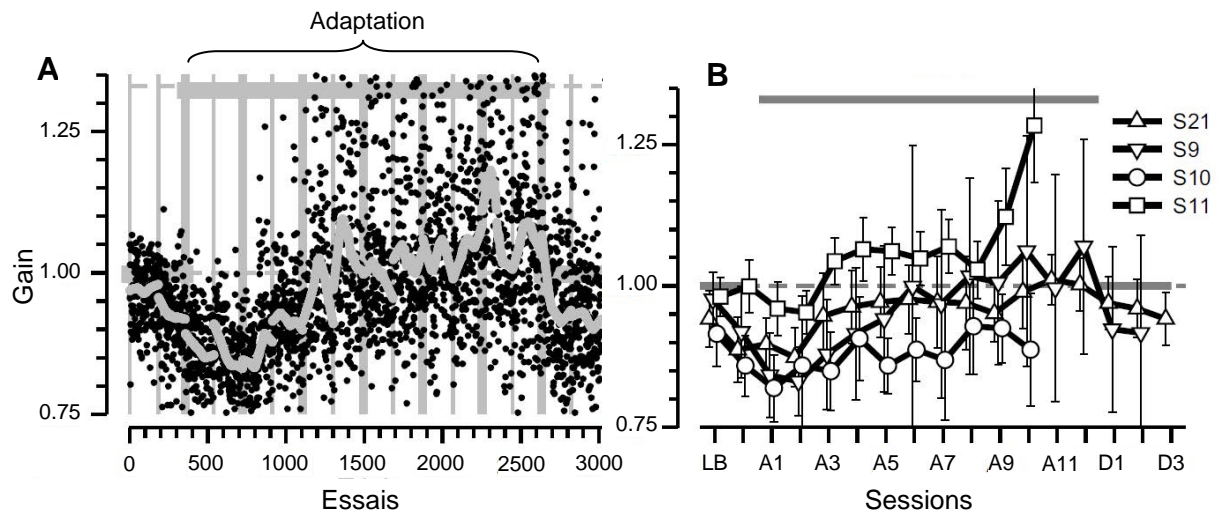


Figure 10. Evolution du gain saccadique au cours de l'expérience 2 (façonnage de l'augmentation du gain avec agent renforçateur auditif). A : gain saccadique observé au fil des essais pour l'individu S9 – les symboles et les lignes sont identiques à ceux de la figure 8. B : gain moyen (et écart-type) calculé sur les 50 derniers essais de chaque session pour les quatre participants.

Tous les participants sauf un (S10) ont vu leur gain saccadique augmenter au cours des sessions d'adaptation renforcée (figure 10 B). En général, les écarts-types des gains étaient faibles, du moins pour les participants S21, S10 et S11 (ils étaient respectivement de 0.06, 0.08 et 0.06 en moyenne, sur les sessions d'adaptation), à l'exception du participant S9 pour qui l'écart-type du gain était de 0.15.

La figure 11 indique les résultats des comparaisons par paires entre les changements de gain obtenus à la fin des différentes conditions expérimentales. Les tests ont effectivement révélé une augmentation significative du gain pour les participants S21, S9 et S11 entre la fin de la ligne de base renforcée et les derniers essais d'adaptation (tests *t* corrigés, $p < 0.05$).

Pour les participants S21 et S9 qui ont effectué des sessions de désadaptation, la modification des contingences de renforcement a également été suivie d'une diminution significative du gain saccadique par rapport au gain observé à la fin de la condition d'adaptation.

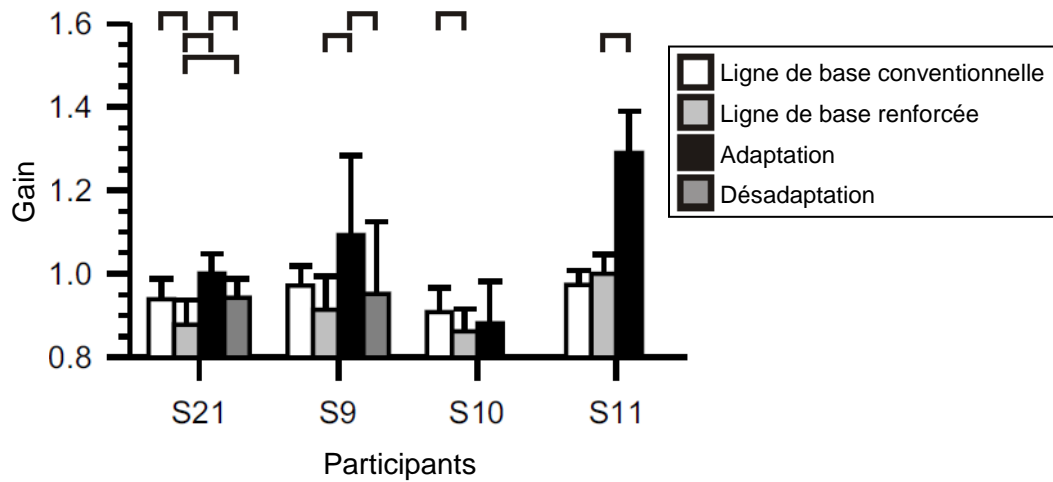


Figure 11. Comparaison des gains moyens (et écarts-types) calculés sur les 50 derniers essais de chaque condition expérimentale pour les quatre participants de l'expérience 2 (façonnage de l'augmentation du gain avec agent renforçateur auditif). Les crochets indiquent les différences significatives obtenues lors des tests *t* corrigés de comparaisons par paires a priori ($p < 0.05$).

La figure 12 permet de comparer les proportions de changements de gain (de l'adaptation par rapport à la ligne de base renforcée) de l'expérience 1 à celles de l'expérience 2. Les procédures de ces expériences se distinguaient seulement par l'amplitude visée par le programme de renforcement percentile (respectivement 9.4° et 18.9°).

Le graphique souligne les différences d'amplitudes induites par les deux types de contingences de renforcement : en effet, tandis que l'amplitude a baissé dans l'expérience 1 (elle a diminué de -23.1% en moyenne), dans l'expérience 2 elle a augmenté ou est restée constante (elle a changé de +16.2% en moyenne). Ces différences entre les deux expériences ont été attestées par les résultats d'un test de Wilcoxon Mann-Whitney effectué sur les quatre premières sessions d'adaptation. Si les amplitudes saccadiques des deux groupes ne différaient pas lors des deux premières sessions, elles ont commencé à diverger lors de la troisième ($p = 0.029$) : les contingences de renforcement différentes ont donc bien induit des adaptations différentes.

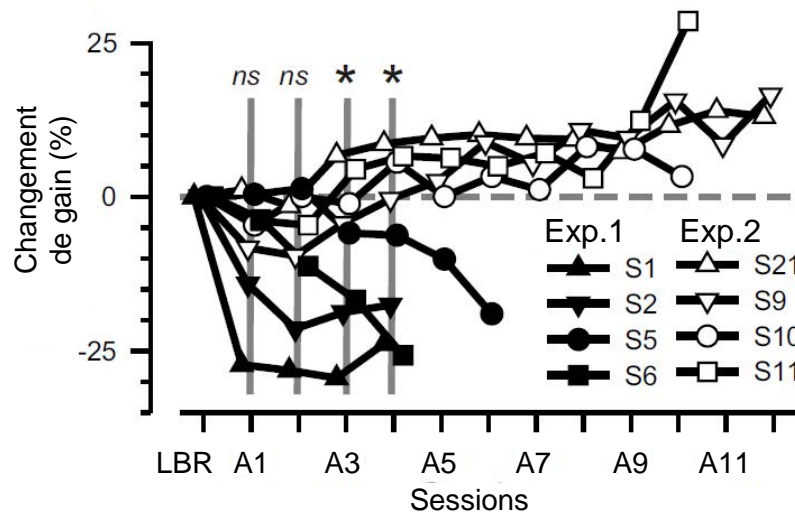


Figure 12. Pourcentage de changement de gain par rapport à la ligne de base renforcée, pour chaque session d'adaptation des expériences 1 (symboles pleins) et 2 (symboles vides), calculé sur les 50 derniers essais. La ligne horizontale en pointillés correspond à une absence de changement, et les lignes verticales indiquent les résultats (ns : non significatifs, astérisque : significatifs) du test de Wilcoxon Mann-Whitney.

2.3. Expérience 3 : façonnage d'une diminution du gain saccadique avec agent renforçateur visuel

Lorsque la cible était éteinte au début des saccades et réaffichée en vision centrale seulement après les saccades dont l'amplitude était de plus en plus courte (jusqu'à une valeur de 9.4°), l'amplitude a diminué également. Le décours temporel de l'adaptation du gain saccadique de l'un des participants de cette expérience (S4) est présenté sur la figure 13 A.

Plusieurs éléments sont intéressants dans ces résultats individuels : tout d'abord, à la fin de la ligne de base renforcée le gain avait déjà commencé à diminuer (le gain moyen des 50 derniers essais de cette condition expérimentale était de 0.82). Ensuite, le gain a baissé progressivement au cours des sessions d'adaptation, pour atteindre une moyenne de 0.67 lors des 50 derniers essais de cette condition expérimentale – cette valeur ayant été observée aux alentours de la deuxième session. Les courbes Lowess montrent qu'au début de chaque session d'adaptation, le gain remontait transitoirement avant d'atteindre un niveau de gain inférieur ou équivalent à celui de la

session précédente. Enfin, le gain est remonté lors de la désadaptation : lors des 50 derniers essais, il était de 0.80 en moyenne.

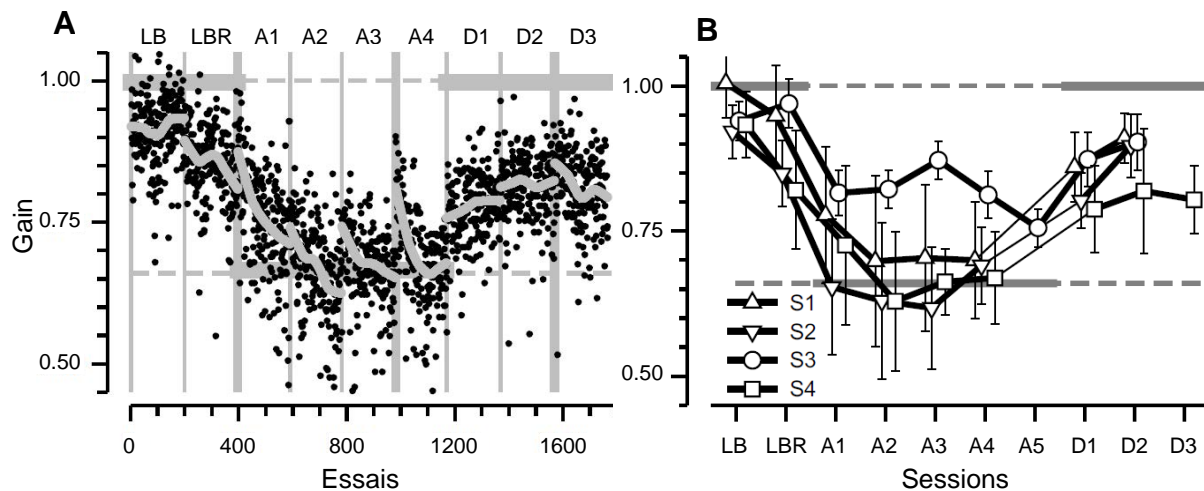


Figure 13. Evolution du gain saccadique au cours de l'expérience 3 (façonnage de la diminution du gain avec la réapparition de la cible en vision centrale en guise d'agent renforçateur). A : gain saccadique observé au fil des essais pour l'individu S4 – les symboles et les lignes sont identiques à ceux de la figure 8. B : gain moyen (et écart-type) calculé sur les 50 derniers essais de chaque session pour les quatre participants.

Les évolutions des changements de gain pour les participants de cette expérience (figure 13 B) sont similaires à ceux de l'expérience 1. Sur les 50 derniers essais d'adaptation, l'amplitude a changé de -21.7% en moyenne par rapport à la ligne de base. Une hausse était également visible lors des sessions de désadaptation.

La figure 14 indique que, selon les résultats aux tests t des comparaisons par paires, la diminution des amplitudes induite par le renforcement lors de la condition d'adaptation était significative pour tous les participants, de même que l'augmentation observée lors de la désadaptation.

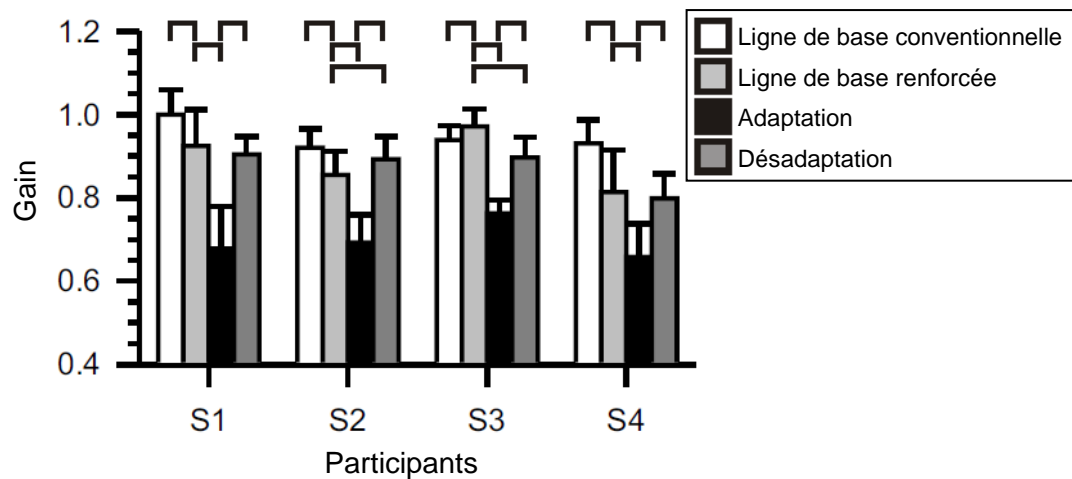


Figure 14. Comparaison des gains moyens (et écarts-types) calculés sur les 50 derniers essais de chaque condition expérimentale pour les quatre participants de l'expérience 3 (façonnage de la diminution du gain avec la réapparition de la cible en vision centrale en guise d'agent renforçateur). Les crochets indiquent les différences significatives obtenues lors des tests *t* corrigés de comparaisons par paires a priori ($p < 0.05$).

2.4. Expérience 4 : façonnage d'une diminution du gain saccadique avec agent renforçateur auditif et stabilisation de l'image de la cible à l'issue de chaque essai

Le fait de stabiliser l'image de la cible sur la fovéa à l'issue de chaque saccade exécutée vers la cible n'a pas empêché le façonnage de pouvoir réduire le gain saccadique dans cette expérience également, où un son était utilisé comme agent renforçateur.

La figure 15 montre que l'amplitude des saccades a diminué chez tous les participants. La proportion moyenne de changement de gain était de -28.5%. Le gain a augmenté lors des sessions de désadaptation – excepté pour le participant S7, dont l'amplitude a continué de diminuer.

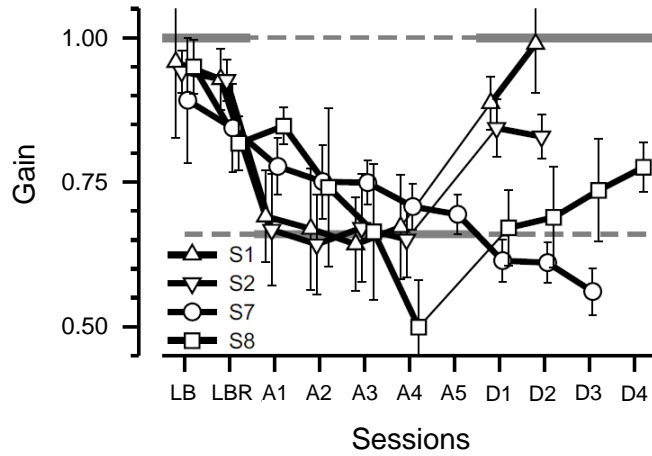


Figure 15. Evolution du gain moyen (et écart-type) calculé sur les 50 derniers essais de chaque session pour les quatre participants de l'expérience 4 (façonnage d'une diminution du gain saccadique avec agent renforçateur auditif et stabilisation de l'image de la cible à l'issue de chaque essai).

Comme dans les expériences 1 et 3, ces observations sont attestées par les résultats des tests t de comparaisons par paires (figure 16). La diminution du gain entre l'adaptation et la désadaptation constatée chez le participant S7 est significative aussi.

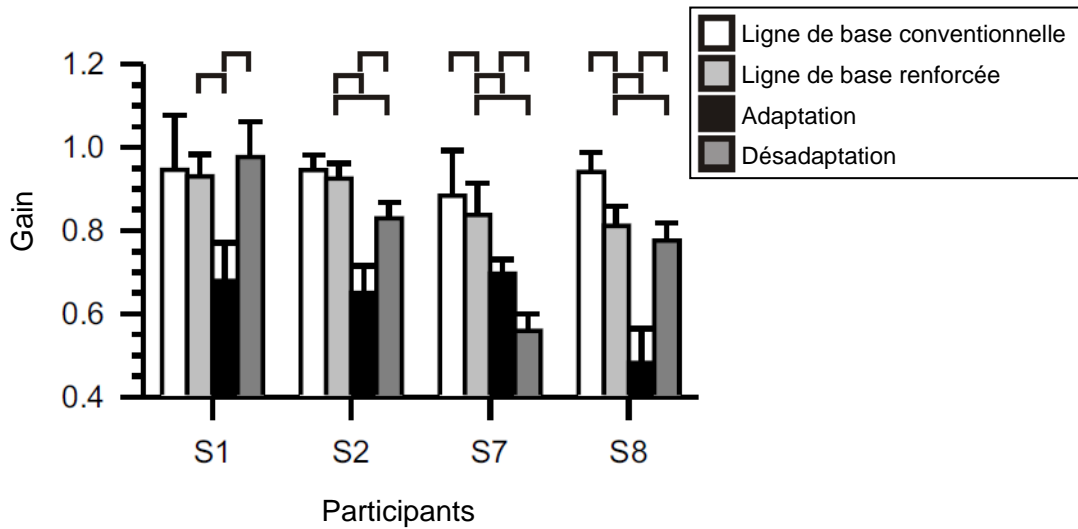


Figure 16. Comparaison des gains moyens (et écarts-types) calculés sur les 50 derniers essais de chaque condition expérimentale pour les quatre participants de l'expérience 4 (façonnage d'une diminution du gain saccadique avec agent renforçateur auditif et stabilisation de l'image de la cible à l'issue de chaque essai). Les crochets indiquent les différences significatives obtenues lors des tests t corrigés de comparaisons par paires a priori ($p < 0.05$).

2.5. Expérience 5 : comparaison entre les adaptations saccadiques induites par la procédure de renforcement et par le paradigme conventionnel

Quatre participants ont passé successivement le paradigme de double saut et celui du façonnage (avec agent renforçateur auditif, l'image de la cible apparaissant stabilisée sur la fovéa à la fin de chaque essai). Cela a permis de comparer les diminutions de gain entraînées par ces deux procédures, les récupérations ainsi que les transferts d'adaptation à des saccades dont l'amplitude ou la position initiale était différente de celle des saccades adaptées.

2.5.1. Comparaison des diminutions de gain induites par les deux procédures

La figure 17 présente l'évolution du gain observée lors des deux adaptations, pour lesquelles le gain cible était 0.66. Pour tous les participants, le gain saccadique a baissé progressivement lors de l'adaptation par façonnage, pour atteindre une moyenne de 0.67. Puis, une semaine après, une nouvelle diminution a été induite par le paradigme conventionnel : le gain moyen était de 0.68 à la fin des sessions d'adaptation. Par rapport à la ligne de base, le gain a été modifié de -22.8% par la procédure de renforcement et de -27.4% par le paradigme de double saut.

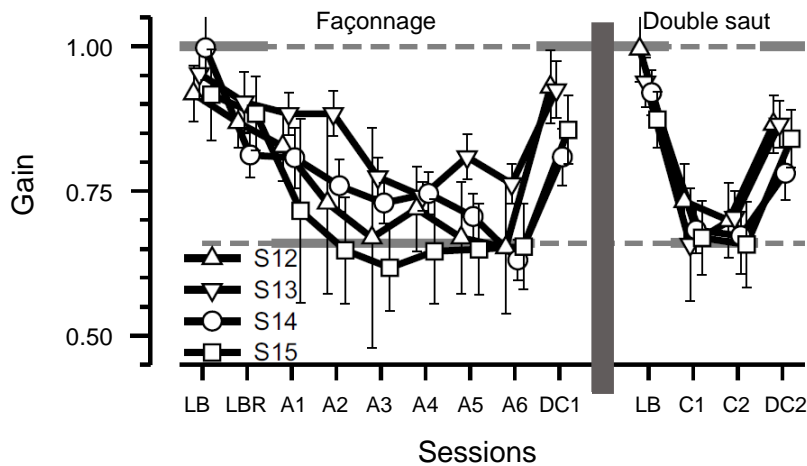


Figure 17. Evolution du gain moyen (et écart-type) calculé sur les 50 derniers essais de chaque session pour les quatre participants des procédures de façonnage et de double saut (expérience 5). Les lignes horizontales indiquent les valeurs cibles, et la ligne verticale matérialise la semaine s'écoulant entre les deux passations ; les lettres indiquent les types de sessions : LB : ligne de base conventionnelle, LBR : ligne de base renforcée, A : adaptation par façonnage, C : adaptation par la procédure conventionnelle et DC : désadaptation conventionnelle.

Par ailleurs, le décours temporel de l'adaptation était plus long et plus variable lors de la procédure de façonnage que lors du paradigme de double saut. En effet, globalement, lors du paradigme conventionnel, l'adaptation saccadique était achevée en une session – autrement dit, la réduction du gain observée correspondait à ce qui était attendu étant donné un saut intra-saccadique négatif de la cible de 33% de son premier déplacement. En revanche, lors des sessions de façonnage, dont le programme de renforcement percentile ciblait un gain de 0.66, le gain est descendu en dessous de 0.7 à la fin de la deuxième session pour le participant S15, à la fin de la troisième pour S12, et de la sixième pour S14.

Les résultats des tests *t* de comparaisons par paires, indiqués sur la figure 18, montrent que les diminutions de gain induites par le façonnage sont significatives pour tous les participants, de manière comparable à ce qui a été observé dans les expériences précédentes. Le gain a été significativement réduit par la procédure conventionnelle aussi. Il a atteint des valeurs plus faibles après la procédure de façonnage qu'après celle du double saut pour un individu (S14), l'inverse a

été observé pour S13, tandis qu'il n'y avait pas de différence pour les deux autres participants (S12 et S15).

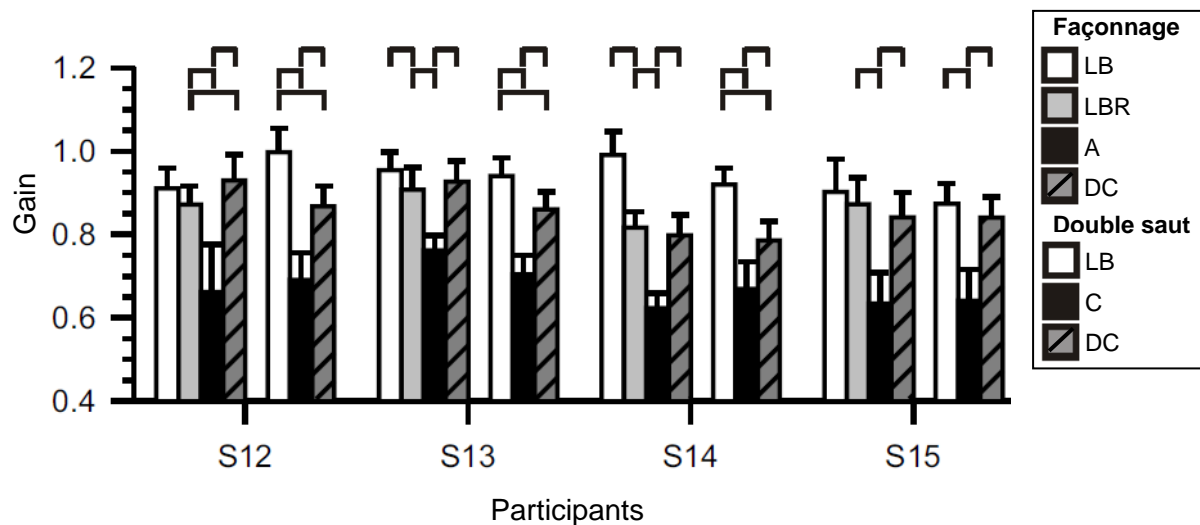


Figure 18. Comparaison des gains moyens (et écarts-types) calculés sur les 50 derniers essais des conditions expérimentales des procédures de façonnage et de double saut pour les quatre participants de l'expérience 5. Les quatre premières barres correspondent à la procédure de façonnage (avec agent renforçateur auditif et stabilisation de l'image de la cible à l'issue de chaque essai) et les trois dernières au paradigme de double saut. LB : ligne de base conventionnelle, LBR : ligne de base renforcée, A : adaptation par façonnage, C : adaptation par la procédure conventionnelle et DC : désadaptation avec cible stationnaire. Les crochets indiquent les différences significatives obtenues lors des tests t corrigés de comparaisons par paires a priori ($p < 0.05$).

Enfin, quel que soit le paradigme utilisé, le gain a augmenté significativement pour tous les participants lors de la désadaptation conventionnelle (dans laquelle l'image de la cible était non stabilisée et la cible restait stationnaire quelle que soit la procédure appliquée). Pour trois personnes sur quatre, le gain est revenu au niveau de ligne de base lors de la désadaptation réalisée après le façonnage. Il était plus élevé après le façonnage qu'après le paradigme conventionnel pour deux participants (S12 et S13). Aucune différence n'a été constatée entre ces deux désadaptations pour les deux autres participants.

2.5.2. Comparaison des désadaptations

Afin de comparer plus précisément les désadaptations réalisées à la fin des deux types de procédures, le gain de chaque essai a été exprimé en fonction du gain moyen calculé lors des 100 premiers essais de la session (des essais d'adaptation), et ce sujet par sujet. Cette proportion était donc calculée de la manière suivante :

$$\frac{\text{gain de l'essai} - \text{gain moyen lors des essais d'adaptation}}{\text{gain moyen lors des essais d'adaptation}} \times 100 \quad (3)$$

Un pourcentage de changement de 0% signifie donc une absence de désadaptation tandis qu'un pourcentage de 100% indique que le gain a doublé par rapport à l'adaptation. Une régression linéaire a ensuite été effectuée sur les 100 premiers essais de chaque désadaptation (la figure 19 A fournit un exemple de cette analyse pour l'un des participants).

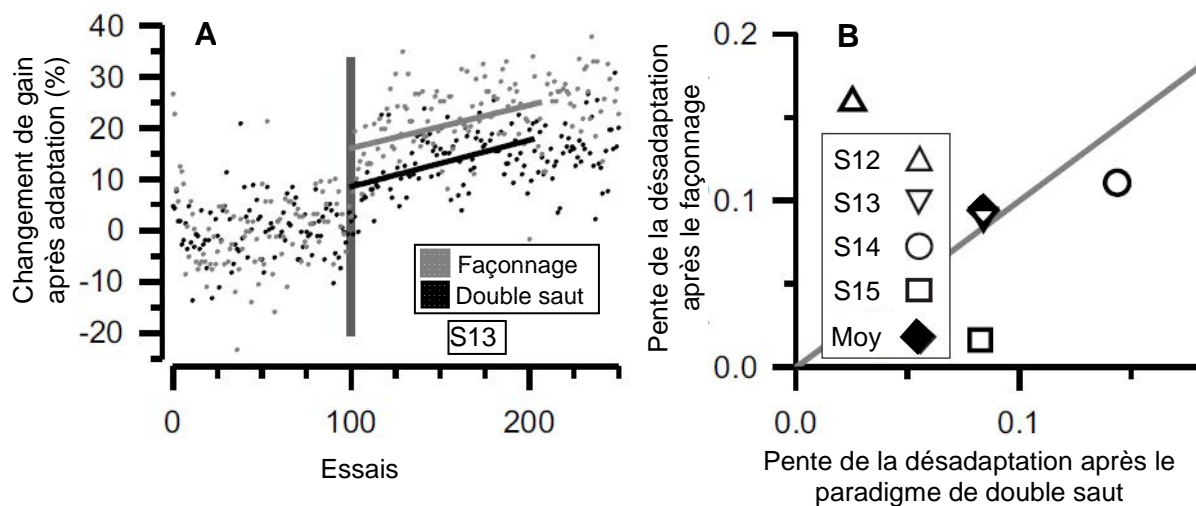


Figure 19. Comparaison des désadaptations observées après les procédures de façonnage et de double saut. A : évolution du pourcentage de changement de gain par rapport à l'adaptation, au fil des essais, pour le participant S13. Un point correspond à un essai. Les droites de régression linéaire ont été calculées pour les 100 premiers essais de la désadaptation. La ligne grise verticale marque le début de la désadaptation. B : Pente de la droite de régression linéaire de la désadaptation après le façonnage en fonction de celle obtenue après le paradigme de double saut, pour chaque participant et en moyenne (losange noir).

La figure 19 B indique les pentes obtenues pour le façonnage en fonction de celles obtenues pour le paradigme conventionnel pour chaque participant. Ces données confirment que la proportion de changement constatée sur ces 100 essais après le façonnage était supérieure à celle observée après le paradigme conventionnel pour le participant S13 mais surtout pour S12. L'inverse a été observé pour S14 et S15.

2.5.3. Comparaison des transferts d'adaptation

Les tests de transfert d'adaptation (transfert « de position » quand la croix de fixation apparaissait à différents endroits, et transfert « d'amplitude » quand la cible apparaissait à différentes excentricités) réalisés à l'issue des sessions de façonnage et de double saut ont révélé que le gain saccadique avait été réduit aussi pour d'autres saccades. Les résultats de ces tests de transfert sont résumés par la figure 20.

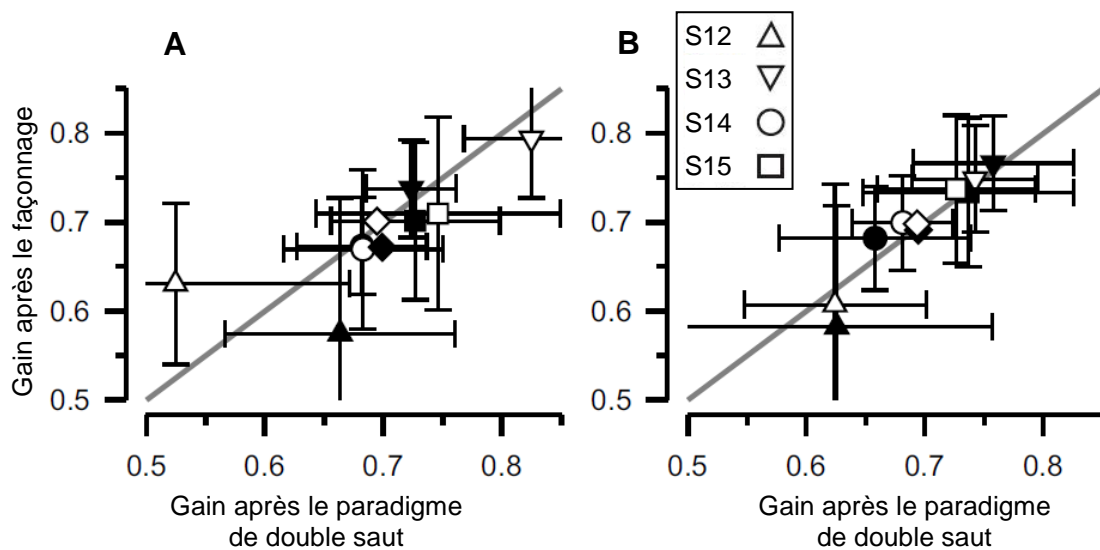


Figure 20. Comparaison entre les transferts d'adaptation obtenus après la procédure de façonnage et ceux observés après la procédure de double saut. A : gains moyens (et écarts-types) des essais de transfert de position (symboles pleins : position initiale de 9.4° ; symboles vides : position initiale de 4.7°). B : gains moyens (et écarts-types) des essais de transfert d'amplitude (symboles pleins : excentricité de la cible de 16.5° ; symboles vides : excentricité de 11.8°).

L'importance du transfert après l'adaptation par le façonnage était similaire à celle observée après le paradigme de double saut. En effet, en ce qui concerne le transfert de position, les gains moyens étaient respectivement de 0.67 et 0.70 lorsque la position initiale de la saccade était 9.4° et de 0.70 et 0.69 lorsque cette position était 4.7°. De même, les gains moyens étaient similaires lors des essais de transfert d'amplitude : ils étaient respectivement de 0.69 et 0.69 pour les cibles apparaissant à une excentricité de 16.5°, et de 0.70 et 0.69 pour une excentricité de 11.8°. Le participant S12 s'est démarqué, avec un gain plus faible après le façonnage pour les essais dont la position initiale était 9.4° (0.57 versus 0.66, $p = 0.01$) et plus élevé après le façonnage pour les essais dont la position initiale était 4.7° (0.62 versus 0.52, $p = 0.001$).

Excepté pour ce participant, quelle que soit la procédure employée, les gains observés lors des essais de transfert étaient tous équivalents ($\pm 5\%$) aux gains obtenus à la fin de l'adaptation. Ce transfert quasiment parfait montre que la diminution du gain induite par les procédures d'adaptation n'était pas l'artéfact d'une stratégie consistant à déplacer l'œil vers une position spécifique.

Enfin, les gains saccadiques moyens des participants qui ont passé uniquement le paradigme conventionnel sont indiqués en bas de la figure 21. La réduction du gain pour chacune des trois personnes (S3, S16 et S17) qui ont passé l'expérience avec un saut intra-saccadique négatif était significative (tests t corrigés, $p < 0.05$). En moyenne, le gain est passé de 0.94 à 0.72, ce qui correspond à un changement de gain de -22.6% (pour un saut intra-saccadique de -33%).

Quant aux participants S18, S19 et S20 qui n'ont passé que le paradigme dans lequel la cible effectuant un saut intra-saccadique positif de 33% de son premier saut, leur gain a significativement augmenté (tests t corrigés, $p < 0.05$). En moyenne, il est passé de 0.90 à 1.02, soit une augmentation de 13%. Notons que ce changement est moins important que la diminution du gain décrite ci-dessus.

Les désadaptations suite à ces deux paradigmes conventionnels étaient elles aussi significatives pour tous les participants (tests t corrigés, $p < 0.05$). Les proportions de changements étaient respectivement de 23.7% après la diminution du gain et de -6.3% après l'augmentation. Cependant, dans les deux cas, le gain n'a pas retrouvé son niveau de base (tests t corrigés, $p < 0.05$).

2.6. Remarques générales sur les résultats des expériences de façonnage

Quel que soit l'agent renforçateur, (agent renforçateur auditif, apparition de la cible en vision centrale, ou son avec stabilisation systématique de l'image de la cible sur la fovéa), les réductions du gain induites par les procédures de façonnage visant une amplitude de 9.4° étaient importantes.

Comme le montre la figure 21, les diminutions étaient similaires : les pourcentages moyens de changement étaient de -23% (agent renforçateur auditif), de -22% (apparition de la cible en vision centrale) et de -28% et -23% (dans les expériences 4 et 5 où l'agent renforçateur était un son et où l'image de la cible était stabilisée sur la fovéa). En outre, le changement de gain obtenu par les procédures de façonnage (-24% en moyenne) était identique à celui obtenu par la procédure conventionnelle (-25% en moyenne).

Suite au changement des contingences de renforcement, à la fin des sessions de désadaptation des expériences 1, 3 et 4, le gain est quasiment revenu à son niveau de base (il a changé de 21.7% en moyenne), excepté pour les participants S6 et S7. Globalement, les augmentations de gain lors de ces désadaptations étaient similaires quelle que soit la procédure de façonnage utilisée.

Quant à l'expérience 2 visant à augmenter le gain par façonnage, la proportion de changement par rapport à la ligne de base était de +16.2% en moyenne, contre 13% via le paradigme de double saut. Le changement des contingences de renforcement, ciblant à nouveau une amplitude de 14.2° , a fait diminuer le gain, qui est alors presque revenu à son niveau de base.

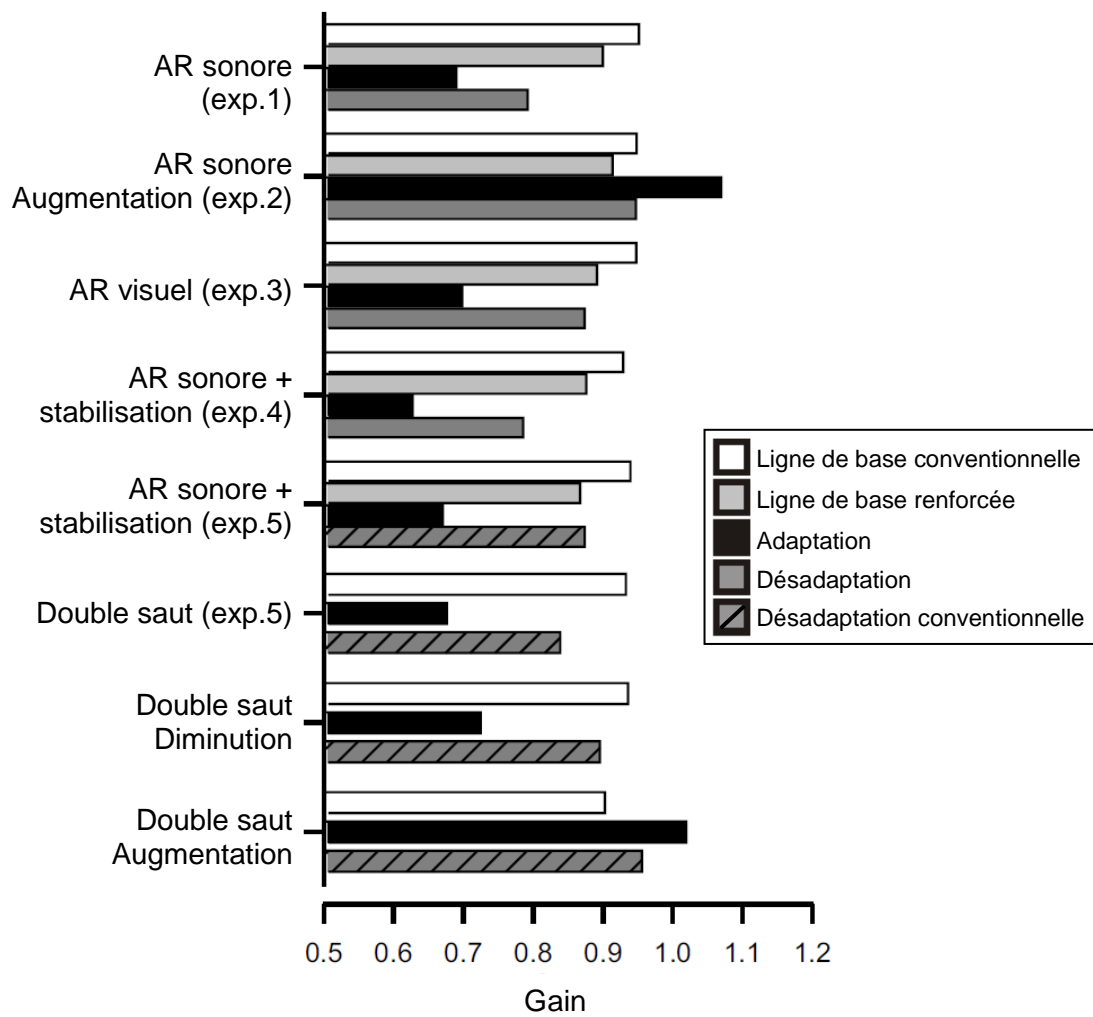


Figure 21. Résumé des résultats des différentes expériences. Les barres représentent les gains moyens calculés à partir des 50 derniers essais de chaque condition expérimentale des différents participants. Les deux barres inférieures correspondent aux résultats des participants qui n'ont passé que le paradigme de double saut (AR = agent renforçateur).

III. Discussion

Ces résultats montrent qu'en l'absence d'erreur de position rétinienne à l'issue des saccades, un agent renforçateur arbitraire (un stimulus sonore ou l'apparition de la cible en vision centrale) peut induire des changements de gain saccadique équivalents à ceux qui résultent d'une procédure de double saut.

Un certain nombre de recherches, détaillées dans la première partie de cette thèse, montrent que l'erreur rétinienne post-saccadique en tant que telle n'est pas suffisante pour contrôler l'adaptation saccadique, mais qu'elle semble devoir être intégrée à un processus de traitement de l'information plus élaboré (ex : Bahcall & Kowler, 2000; Collins et al., 2007; Collins & Wallman, 2011; Havermann & Lappe, 2010; Herman et al., 2009; Madelain et al., 2010; Wong & Shelhamer, 2011). Ainsi, les théories classiques du contrôle moteur postulent l'existence de mécanismes internes qui utiliseraient cette erreur rétinienne pour générer des signaux spécifiques permettant de calibrer un gain interne.

En revanche, nos résultats plaident en faveur d'un mécanisme général d'apprentissage opérant qui repose sur les conséquences *fonctionnelles* des mouvements, c'est-à-dire des conséquences définies essentiellement par leur capacité (démontrée expérimentalement) à contrôler des changements comportementaux. Contrairement à l'erreur rétinienne post-saccadique et au signal qui en découlerait, ces conséquences ne sont donc pas nécessairement de nature visuelle et ne sont pas non plus définies par un signe.

Ces données peuvent amener à faire l'hypothèse qu'au sein du système oculo-moteur aussi se produirait une sélection des mouvements qui sont suivis de conséquences bénéfiques. Lors des expériences d'adaptation saccadique, les saccades qui arrivent de plus en plus près de la cible à l'issue de son second saut, donc suivies de la vision nette de la cible, seraient reproduites, contrairement à ceux qui arrivent trop loin de cette cible. En d'autres termes, l'ajustement du gain observé semble pouvoir s'expliquer par un processus de sélection comportementale (Donahoe, 2003; Donahoe et al., 1993).

Cette discussion résume tout d'abord les éléments qui montrent que les contingences de renforcement sont bien à l'origine des changements de gain saccadique obtenus. Puis elle présente les principales différences et enfin les similarités entre les adaptations saccadiques induites par le paradigme conventionnel de double saut et par la procédure de façonnage.

Plusieurs hypothèses explicatives sont envisagées. La littérature permet d'avancer un certain nombre d'hypothèses cognitives, ce qui s'explique en grande partie par l'abondance des recherches sur l'adaptation saccadique réalisées par les théories classiques du contrôle moteur. Cependant, nous verrons que certains arguments peuvent être interprétés à la lumière de l'analyse expérimentale du comportement.

3.1. Les contingences de renforcement sont responsables des changements de gain saccadique

Plusieurs raisons amènent à considérer les contingences de renforcement comme responsables des changements d'amplitude. En premier lieu, le gain est effectivement modifié en *fonction* de ces contingences – il diminue lors des sessions d'adaptation dans les expériences où les procédures de renforcement percentile visent une diminution du gain, et il augmente ou ne change pas dans les expériences où une augmentation du gain est renforcée.

En second lieu, la seule différence qui intervient entre les sessions d'adaptation et de désadaptation dans la procédure de façonnage réside dans le changement de la valeur cible. Ce changement a entraîné des modifications de gain subséquentes pour quasiment tous les participants, comme le résume la figure 22 A. Sur ce graphique, les diminutions du gain lors de l'adaptation suivies par des augmentations lors de la désadaptation (expériences 1, 3, 4 et 5) sont censées apparaître dans le quadrant en haut à gauche. Quant aux augmentations du gain lors de l'adaptation suivies par des diminutions lors de la désadaptation (expérience 2), elle sont censées apparaître dans le quadrant en bas à droite. C'est effectivement ce qui apparaît sur cette figure.

A l'exception de deux participants, pour tous les autres les changements de gain sont similaires à ceux obtenus avec le paradigme conventionnel de double saut (symboles gris). Le type d'agent renforçateur utilisé, lui, n'a pas d'effet, qu'il soit auditif ou visuel (figure 21). Ces éléments montrent que les changements de gain étaient effectivement contrôlés par les conséquences fonctionnelles des saccades.

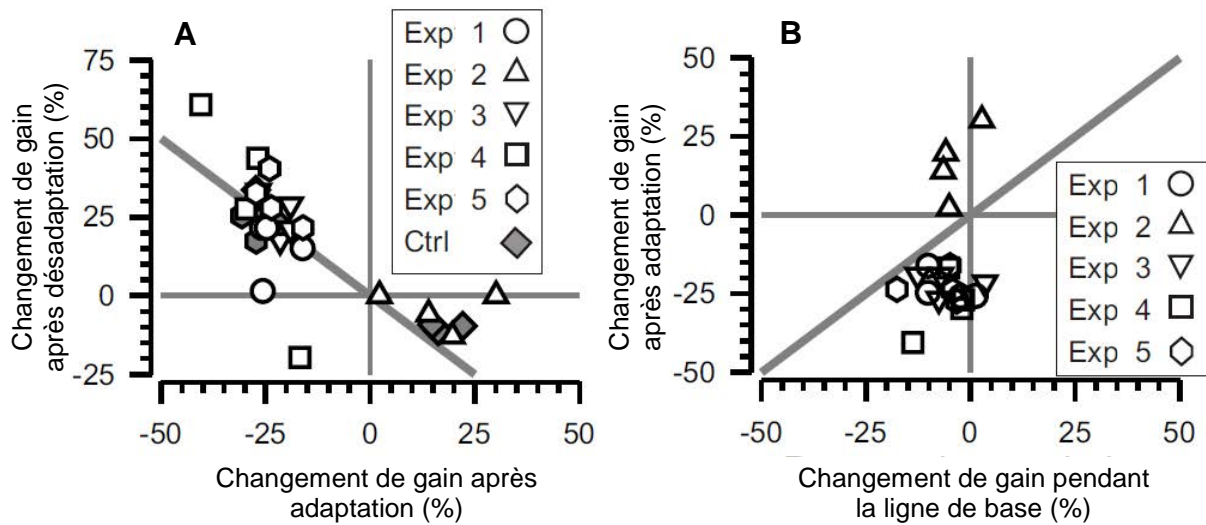


Figure 22. Relations entre les pourcentages moyens de changements de gain obtenus lors des différentes expériences. A : proportion de changement de gain moyen à la fin de la désadaptation en fonction de celui de l'adaptation (ctrl : contrôle, c'est-à-dire paradigme conventionnel). B : proportion de changement de gain moyen à la fin de l'adaptation en fonction de celui de la ligne de base renforcée.

La diminution du gain souvent observée lors de la ligne de base renforcée pourrait remettre en cause cette interprétation des résultats. Cette légère diminution (entre la première ligne de base à la seconde, le gain passe de 0.94 à 0.89 en moyenne) est observée dans 60% des cas, alors que le programme de renforcement percentile cible un gain de 1. Si les sessions de ligne de base renforcée avaient été poursuivies plus longtemps, la diminution du gain aurait peut-être pu s'accroître, indépendamment des contingences de renforcement – ce qu'il aurait éventuellement fallu tester expérimentalement, avec un groupe contrôle supplémentaire.

Cette diminution du gain pourrait s'expliquer par le fait qu'un mécanisme maintiendrait activement une certaine hypométrie, comme le suggèrent les recherches dans lesquelles une erreur constante est imposée au système oculo-moteur (Havermann & Lappe, 2010; Henson, 1978; Wong & Shelhamer, 2011).

De même, les façonnages de nos expériences 3 à 5 pourraient aussi amplifier cette tendance à l'hypométrie. En effet, le procédé de stabilisation de l'image de la cible pourrait simuler des situations où les saccades paraissent plus grandes que celles qui sont émises habituellement, et donc induire une diminution du gain. Quant à la ligne de base de l'expérience 1, lorsque la cible

est éteinte au début des saccades, aucune information ne viendrait contrarier cette tendance à l'hypométrie.

Toutefois, cette diminution lors de la ligne de base ne peut pas expliquer la réduction du gain induite par les procédures de façonnage pour au moins deux raisons. D'une part, notamment dans les expériences 3 à 5, les sessions de désadaptation induisent une augmentation du gain même si l'image de la cible est encore stabilisée sur la fovéa. Cela n'est attribuable qu'aux contingences de renforcement. D'autre part, la diminution du gain observée lors de l'adaptation n'est pas corrélée à celle qui a pu se produire lors de la ligne de base, comme le montre la figure 22 B.

3.2. Différences entre les changements de gain résultant des procédures de façonnage et ceux produits par le paradigme de double saut

3.2.1. Des pics de vitesse saccadique plus élevés lors du façonnage

Les différences entre les effets des façonnages et ceux du paradigme de double saut pourraient se refléter au niveau des propriétés des saccades. Par conséquent, les dynamiques des mouvements ont été comparées. Pour chaque essai, le pic de vitesse a été exprimé en fonction de l'amplitude saccadique selon l'équation proposée par Lebedev, Van Gelder et Tsui (1996) :

$$\text{Pic de Vitesse} = a\sqrt{\text{amplitude}} . \quad (4)$$

Le paramètre a est donc proportionnel au pic de vitesse.

La figure 23 A présente les résultats de l'expérience 5, dans laquelle les participants ont effectué les deux types d'adaptations. Les valeurs de ce paramètre obtenues pour la dernière session d'adaptation du façonnage et celle du paradigme de double saut ne diffèrent pas significativement selon la procédure utilisée (test t apparié, $p = 0.15$).

En ce qui concerne les résultats du paradigme conventionnel, lors de cette expérience 5 (symboles pleins sur la figure 23 A), les valeurs moyennes de ce paramètre ne diffèrent pas non plus entre la ligne de base et la dernière session d'adaptation (les paramètres a moyens sont respectivement de 112 et 114, test t apparié, $p = 0.11$). Quant aux participants de l'expérience 5 qui n'ont passé que le paradigme de double saut (figure 23 B, losanges gris), le paramètre a vaut

105 pour les saccades de la ligne de base comme celles de la fin de l'adaptation (test t apparié, $p = 0.97$).

Cependant, pour les résultats obtenus via la procédure de façonnage, une augmentation est constatée (test t apparié, $p = 0.02$) entre la ligne de base ($a = 108$ en moyenne) et la dernière session d'adaptation ($a = 125$ en moyenne) lors de l'expérience 5 (figure 23 A, symboles vides). De même, lors des autres expériences de façonnage, le paramètre a moyen des dernières sessions d'apprentissage ($a = 112$) est plus élevé que celui de la ligne de base ($a = 105$, test t apparié, $p = 0.03$).

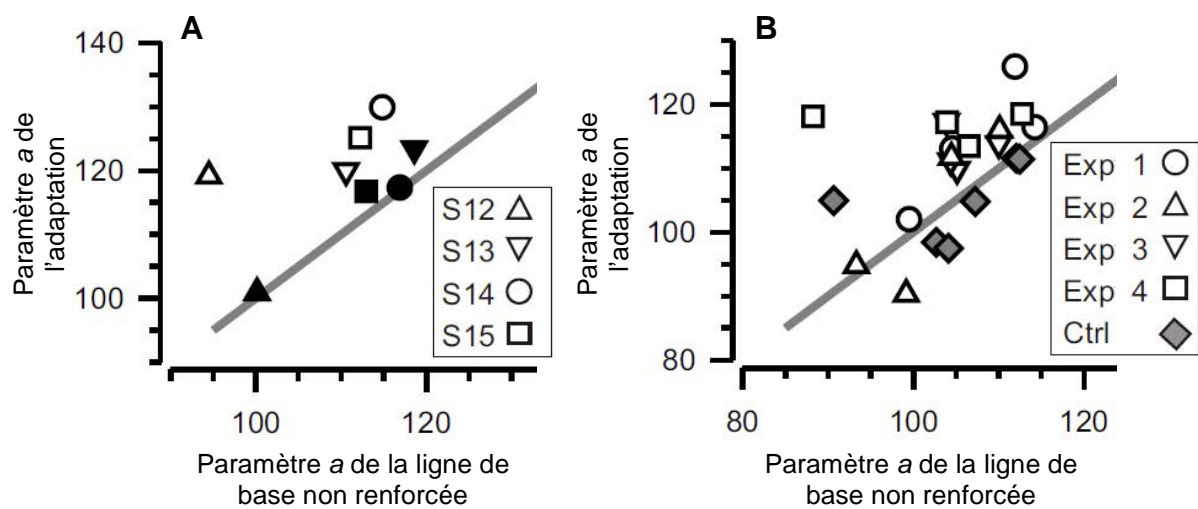


Figure 23. Comparaison des *main sequences* des saccades de ligne de base non renforcée et des différentes procédures d'adaptation. A : paramètre a de chaque participant (S12, S13, S14 et S15) de l'expérience 5. Symboles pleins : paradigme conventionnel ; symboles vides : procédure de façonnage. B : paramètre a des expériences 1 à 4 et des participants qui n'ont effectué que la procédure conventionnelle lors de l'expérience 5 (S3, S16, S17, S18, S19 et S20).

L'étude récente d'Ethier et al. (2008), dans laquelle l'influence de la fatigue est contrôlée, se base sur l'observation que la *main sequence* est modifiée lors du paradigme de double saut visant à réduire le gain saccadique pour inférer l'existence des mécanismes à l'origine des changements de gain – notons qu'à ce jour, il n'existe toutefois pas de consensus au sujet de la modification de la *main sequence* par la procédure d'adaptation (Pélisson et al., 2010).

En revanche, nos résultats ne permettent pas de conclure sur l'éventuelle implication d'un mécanisme spécifiquement lié à la modification du gain saccadique par les contingences de renforcement. En effet, des études semblent indiquer que le renforcement différentiel des saccades, en dehors de tout processus d'adaptation saccadique, peut entraîner la modification de la *main sequence*, que ce renforcement soit contingent à la vitesse des saccades (Madelain & Paeye, 2011) ou non. Ainsi, chez l'être humain, une augmentation de la vitesse a été observée alors que les conséquences des saccades sont contingentes à leur direction (Xu-Wilson et al., 2009) ou à leur latence (Montagnini & Chelazzi, 2005). On ne peut pas exclure un éventuel effet lié à une différence d'arousal. En effet, aucun « bravo » ni point n'était dispensé dans le paradigme de double saut, contrairement aux expériences de façonnage – où le pic de vitesse est plus élevé. Or la simple présentation de stimulus susceptibles de fonctionner comme des agents renforçateurs augmente l'intensité des réponses (Killeen et al., 1978).

3.2.2. Des changements de gain induits par le façonnage plus lents que ceux résultant du paradigme de double saut ?

L'existence d'une variabilité inter et intra-individuelle dans les résultats des études portant sur l'adaptation saccadique a déjà été évoquée précédemment. Celle-ci a été relevée chez l'être humain (Albano & King, 1989; Frens & van Opstal, 1994; Hopp & Fuchs, 2004; Pélisson et al., 2010; van Beers, 2007) et chez le singe (Robinson et al., 2006; Straube et al., 1997). Cette variabilité est en partie liée à la pléiade des procédures et des équipements qu'emploient les différentes études.

Cependant, la variabilité semble encore plus élevée pour les résultats de nos procédures de façonnage – a fortiori, ces données proviennent de notre seule étude. En fait, cette variabilité concerne surtout le décours temporel des adaptations. En effet, chez les participants qui ont passé les deux types de procédures d'adaptation, l'ampleur des changements est similaire (en moyenne, 24% et 25% de diminution ont été obtenus par le façonnage et par la procédure conventionnelle respectivement, ce qui correspond à des gains de 0.67 et 0.70, comme l'illustre la figure 21). Mais pour certains participants, le gain saccadique s'adapte aussi rapidement avec le façonnage qu'avec la procédure conventionnelle, tandis que pour d'autres participants, les changements sont beaucoup plus lents dans le premier cas (voir figure 17).

Pour comparer plus précisément les décours temporels des adaptations induites par les deux types de paradigmes chez les mêmes individus, la moyenne des gains de chaque essai a été calculée à partir des résultats des quatre participants (S12, S13, S14 et S15) de l'expérience 5, en fonction du type de procédure. Ces évolutions ont été résumées par des fonctions exponentielles du second degré (figure 24 A). La constante de temps (c'est-à-dire le nombre d'essais nécessaires pour obtenir 63% de la valeur de l'asymptote) est estimée à 183 essais pour l'adaptation par façonnage et 49 essais pour le paradigme conventionnel. A première vue, les effets du façonnage s'avèrent donc en effet plus lents que ceux de la procédure de double saut.

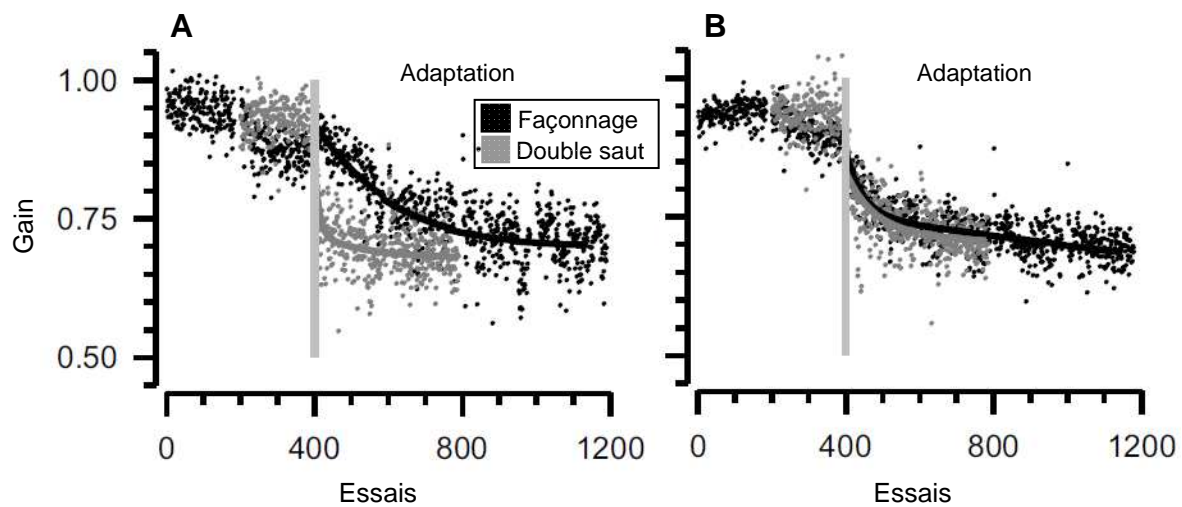


Figure 24. Comparaison des décours temporels des changements du gain saccadique (moyenné sur les gains des différents participants pour chaque essai) au cours des procédures de façonnage et de double saut. Les courbes représentent les fonctions exponentielles correspondantes. A : pour les quatre participants (S12, S13, S14 et S15) de l'expérience 5. B : comparaison des données du façonnage (en noir) à partir des résultats des expériences 1, 3 et 4 et des données des participants qui n'ont passé que le paradigme de double saut (S3, S16 et S17, en gris).

Néanmoins, quand on compare les données obtenues dans les expériences 1, 3 et 4 avec celles des participants qui n'ont passé que le paradigme de double saut (S3, S16 et S17), les deux courbes sont assez semblables et atteignent des asymptotes similaires, avec des constantes de temps proches (108 contre 86 essais, respectivement, figure 24 B).

Cette différence entre les répliques du paradigme conventionnel, soulignée par cette figure 24, pourrait relever de la facilitation de l'adaptation qui se produit quand les apprentissages sont

répétés – c'est ce que suggèrent les études de Kojima et al. (2004), Robinson et al. (2006) ou Körding et al. (2007) chez le singe. Dans notre expérience 5, avant d'être modifié par le paradigme de double saut, le gain saccadique avait déjà été adapté par les procédures de façonnage.

Quant à l'éventuelle lenteur du façonnage par rapport au paradigme de double saut, plusieurs hypothèses explicatives pourraient être avancées.

La première porte sur l'existence éventuelle de plusieurs mécanismes responsables de l'adaptation saccadique. Ainsi, aux mécanismes de réévaluation des coordonnées de la cible et de modification des commandes motrices invoqués par des auteurs comme Chen-Harris et al. (2008), Ethier et al. (2008) ou Semmlow et al. (1989), pourrait s'en ajouter un autre qui reposerait sur le renforcement. Les mécanismes agissant selon des échelles de temps différentes, il serait possible que celui qui est lié au renforcement soit plus lent que les autres.

Cette – hypothétique – conjonction de mécanismes pourrait expliquer aussi un certain nombre de différences entre les conditions expérimentales. Par exemple dans l'expérience 3 où l'agent renforçateur correspond à la seule apparition de l'image de la cible stabilisée, l'adaptation pourrait impliquer le mécanisme de comparaison de l'erreur réelle et de l'erreur prédite, agissant seul ou de concert avec le mécanisme de renforcement et favorisant une diminution du gain, tandis que lors de la désadaptation le mécanisme de renforcement serait ralenti par l'intervention de ce mécanisme. Toutefois cette explication en termes de mécanismes agissant conjointement reste une conjecture.

La seconde hypothèse envisage ce phénomène en termes d'histoire de renforcement. Dans les procédures de façonnage, le système oculo-moteur se trouve confronté à des situations inhabituelles, notamment dans les expériences 1 et 2 où la cible disparaît pendant les saccades. En effet, les conséquences auditives ne sont pas celles sur lesquelles le système visuel se base depuis la vingtaine voire trentaine d'années de vie des participants. Or l'histoire de renforcement est un facteur important qui pèse inévitablement sur l'influence des contingences actuelles, lorsqu'une expérience (ou une intervention comportementaliste à visée éducative) en applique de nouvelles (Lattal & Neef, 1996). Cette influence pourrait expliquer la lenteur des programmes de façonnage par rapport au paradigme de double saut.

3.3. Similarités entre les changements de gain résultant des deux types de procédures

Cette augmentation du pic de vitesse des saccades ainsi que la lenteur des modifications du gain observées lors des procédures de façonnage sont les principaux éléments qui distinguent les effets de cette procédure opérante de ceux résultant du paradigme de double saut. L'examen d'autres caractéristiques, comme le transfert d'adaptation, les latences, la différence entre l'augmentation et la réduction du gain ainsi que la rétention de l'adaptation, montre dans quelle mesure ces adaptations sont comparables. Ces éléments indiquent aussi que l'utilisation éventuelle de stratégies cognitives ne peut pas rendre compte des adaptations saccadiques induites par les deux procédures.

3.3.1. Quelle que soit la procédure utilisée, l'adaptation saccadique ne semble pas résulter de stratégies cognitives

De manière générale, les stratégies cognitives se définissent, indépendamment de leur degré d'élaboration et du fait qu'elles soient conscientes ou non, comme des processus qui sont impliqués dans le traitement visuel, la sélection de cibles ainsi que la programmation motrice, et qui dépendent des propriétés des stimulus et de l'information stockée en mémoire (Glimcher, 2003).

L'utilisation de telles stratégies est une question récurrente dans les études sur l'adaptation saccadique qui utilisent le paradigme de double saut (Collins & Doré-Mazars, 2009; Hopp & Fuchs, 2004).

a. Des changements de gains progressifs

Le principal argument que ces études avancent contre l'utilisation de telles stratégies réside dans le caractère progressif des changements de gain observés. C'est le cas non seulement pour les adaptations induites par notre paradigme de double saut, mais aussi pour celles résultant des procédures de façonnage. A fortiori, les changements de gain sont plus lents lors du façonnage que lors du paradigme de double saut, comme cela vient d'être exposé.

Dans le même ordre d'idée, le retour graduel du gain à son niveau de base lors des désadaptations réfute l'hypothèse de l'utilisation de stratégies cognitives : si ces dernières étaient utilisées, l'adaptation devrait disparaître dès la fin des essais comportant des doubles sauts. Or les effets de ce type de procédure ne disparaissent pas en un seul essai (Alahyane & Pélisson, 2005; Collins & Doré-Mazars, 2009; Hopp & Fuchs, 2004; Pélisson et al., 2010).

De même que les adaptations, les désadaptations produites par le façonnage sont progressives. En particulier, dans les sessions de désadaptation de l'expérience 5, la cible n'effectue pas de second saut mais reste visible à l'issue des saccades, une situation dans laquelle l'utilisation de stratégies se traduirait probablement par un retour extrêmement rapide à la valeur du gain initial. Or les désadaptations ne diffèrent pas selon la procédure appliquée au préalable (figure 19).

Ce caractère progressif des changements d'amplitude est observé en dépit de la prédictibilité de la situation. En effet, lors de ces expériences, la croix de fixation apparaît toujours à la même position, et l'amplitude du premier saut de la cible est toujours la même. Dans des conditions aussi invariables, la meilleure solution consisterait à découvrir dans l'écran la localisation spatiale qui permet d'obtenir le plus de points possible, et de reproduire les saccades dont l'amplitude correspond à l'excentricité de cette localisation. Si les participants utilisaient une telle stratégie consciente pour maximiser les récompenses, on peut espérer qu'ils soient capables de découvrir cette localisation en moins de quelque 1000 essais réalisés sur plusieurs jours.

b. Un transfert supérieur à 97% en moyenne

Le fait que l'adaptation induite par le façonnage se transfère de manière presque parfaite aux saccades n'ayant pas fait l'objet de contingences de renforcement semble aussi exclure l'utilisation de stratégies – 97% et 98% de transfert ont été obtenus lors des tests de transfert de position et d'amplitude, respectivement. De surcroît ces pourcentages sont similaires à la proportion de transfert qui résulte de la procédure conventionnelle (figure 20).

Ils sont en outre cohérents avec ceux de Frens et van Opstal (1994) ou de Semmlow et al. (1989) par exemple, ou avec ceux de Noto, Watanabe et Fuchs (1999) ou Wallman et Fuchs (1998) en ce qui concerne le singe. Ces auteurs ont montré que l'adaptation obtenue avec le paradigme de double saut se transfère sinon complètement, du moins en grande partie, à des saccades de même vecteur ou de vecteur proche, indépendamment de la position initiale des yeux.

c. Des changements de latences non significatifs ou isolés

Les latences saccadiques sont souvent considérées comme le reflet de diverses stratégies cognitives, relatives par exemple à la prise de décision (Glimcher, 2003; Reddi & Carpenter, 2000; Stritzke et al., 2009). Auparavant, certains auteurs, comme Albano et King (1989) ou Frens et van Opstal (1994), n'observant pas de différence entre les latences des saccades adaptées et non adaptées, en avaient aussi déduit l'absence de l'implication de telles stratégies dans l'adaptation saccadique.

De même, la figure 25 indique que le paradigme de double saut n'entraîne globalement pas de différence entre les latences de la ligne de base non renforcée et celles de la dernière session d'adaptation conventionnelle. La figure 25 A permet de comparer les moyennes des latences saccadiques des participants qui ont effectué les deux types de procédures (S12, S13, S14 et S15) dans l'expérience 5. La figure 25 B trace les latences moyennes observées lors de toutes les autres expériences de façonnage ainsi que celles des participants qui n'ont effectué que la procédure conventionnelle lors de l'expérience 5 (S3, S16, S17, S18, S19 et S20). Entre la ligne de base et les dernières sessions du paradigme conventionnel, la latence moyenne passe de 178 à 190 ms (figure 25 A, symboles pleins) et de 189 à 191 ms (figure 25 B, losanges gris), différences qui ne sont pas significatives (test *t* apparié, $p = 0.22$ et $p > 0.47$).

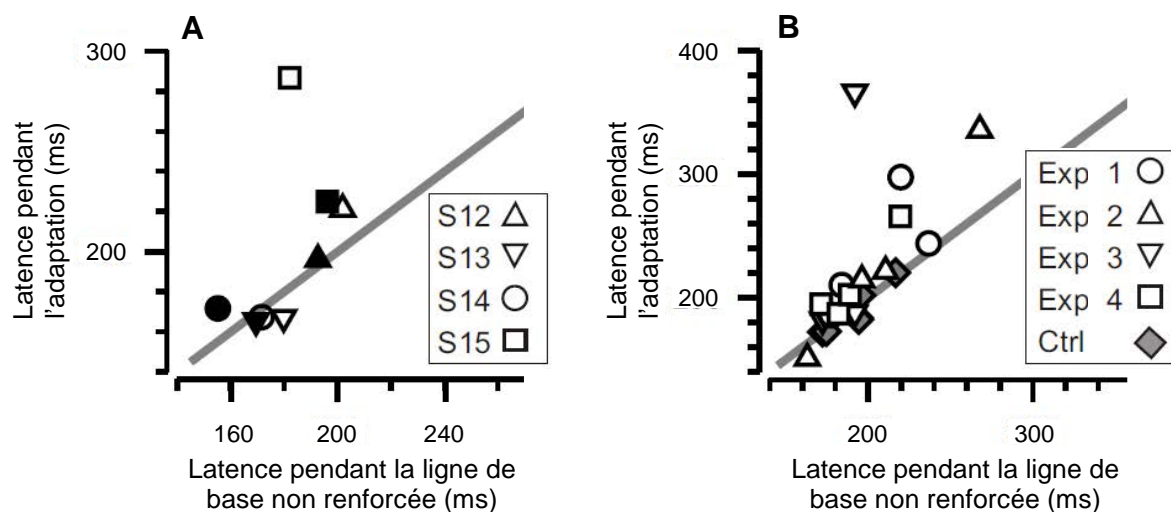


Figure 25. Latence moyenne pendant l'adaptation en fonction de la latence moyenne pendant la ligne de base non renforcée. A : pour les quatre participants (S12, S13, S14 et S15) de l'expérience 5. Symboles pleins : paradigme conventionnel ; symboles vides : façonnage. B : pour les expériences 1 à 4 et les participants qui n'ont effectué que la procédure conventionnelle lors de l'expérience 5 (S3, S16, S17, S18, S19 et S20, losanges gris).

Entre la ligne de base et la dernière session des façonnages, une légère augmentation des latences a lieu, mais elle n'est visible que dans les résultats de la figure 25 B (la latence moyenne passe alors de 197 à 227 ms, test t apparié, $p = 0.02$). Cette augmentation est inexplicable car elle n'est visible que de manière isolée et pour des participants ayant passé des expériences différentes.

Les études sur le renforcement des mouvements oculaires chez le singe soulignent que les saccades émises vers une cible associée à une récompense ont tendance à avoir des latences plus courtes que des saccades émises dans des directions ne signalant pas l'arrivée immédiate de cette récompense (ex : Hikosaka et al., 2006; Lauwereyns et al., 2002; Nakamura & Hikosaka, 2006; Sohn & Lee, 2006; Takikawa et al., 2002; Watanabe et al., 2003). Cela pourrait refléter des contingences de renforcement apprises antérieurement. Chez l'être humain, Madelain et al. (2007) montrent que le renforcement contingent aux latences longues permet de contrôler les distributions de ces dernières. Une explication possible de l'augmentation des latences dans nos expériences de façonnage pourrait résider dans le renforcement accidentel de latences longues, sans que cela soit nécessairement lié à d'éventuelles stratégies impliquées dans la modification du gain saccadique.

Pour les participants dont les latences sont représentées sur la figure 25 A (symboles vides), cette différence – intra-individuelle – n'est pas significative (les latences moyennes passant de 183 lors de la ligne de base à 210 ms à la fin de l'adaptation, test t apparié, $p = 0.39$). Pour eux, la latence moyenne des saccades de la dernière session du façonnage (210 ms) n'est pas différente non plus de celle de la procédure de double saut (190 ms, test t apparié, $p = 0.25$).

En résumé, l'étude des latences ne permet pas de mettre en évidence le fait que les procédures de façonnage impliqueraient des stratégies particulières par rapport au paradigme de double saut.

d. Une augmentation du gain différente de la diminution

Un dernier argument allant à l'encontre des stratégies consiste à comparer les adaptations induites par des sauts intra-saccadiques négatifs à celles résultant de sauts positifs : si une quelconque stratégie était à l'œuvre, alors il n'y aurait pas de raison que les diminutions du gain soient plus rapides que les augmentations.

Or la différence entre les adaptations qui visent à augmenter le gain saccadique et celles qui visent à le diminuer, les premières étant plus rapides et/ou plus importantes que les secondes, est

largement attestée par la littérature sur les recherches qui emploient le paradigme de double saut, chez l'homme (Albano & King, 1989; Bahcall & Kowler, 2000; Collins & Doré-Mazars, 2009; Ethier et al., 2008; Hopp & Fuchs, 2004; Miller et al., 1981; Péliisson et al., 2010), comme chez le singe (Robinson et al., 2003; Straube et al., 1997).

Cette différence se retrouve dans les adaptations induites par la procédure opérante de façonnage. En effet, la valeur absolue du changement de gain observé lors de l'expérience 1 est plus importante que celle de l'expérience 2 (toutes deux utilisant un agent renforçateur auditif et faisant disparaître la cible au début des saccades, cf. figure 12). De même, la désadaptation est incomplète pour cinq participants sur douze, et le gain ne remonte pas du tout pour deux autres participants – toutefois, si les sessions de désadaptations avaient duré plus longtemps, le gain aurait éventuellement pu être augmenté davantage.

Certains auteurs, dont les études chez l'homme (Chen-Harris et al., 2008; Ethier et al., 2008; Semmlow et al., 1989) ou chez le singe (Kojima et al., 2004; Robinson et al., 2006; Körding et al., 2007) ont été présentées dans la première partie de cette thèse, expliquent cette différence entre augmentation et diminution du gain par la coexistence de plusieurs mécanismes internes sous-jacents. Ces mécanismes évolueraient selon des échelles de temps qui leur seraient propres.

Quelles que soient les hypothèses relatives à ces éventuels mécanismes internes, le fait que la diminution du gain soit plus rapide et plus importante que l'augmentation pourrait s'expliquer par le fait que le système oculo-moteur maintienne activement un certain niveau d'hypométrie (Harris, 1995; Havermann & Lappe, 2010; Henson, 1978; Miller et al., 1981; Wong & Shelhamer, 2011). Toutefois, à ce jour, cette différence entre les deux adaptations n'a pas reçu d'explication consensuelle.

3.3.2. Rétention de l'adaptation

Lors des expériences de façonnage d'une diminution de l'amplitude saccadique, les changements de gain étaient en général maintenus d'une session à l'autre, et même d'un jour à l'autre. En effet, l'amplitude moyenne calculée sur les 25 premiers essais de chaque session diminue progressivement au fil des sessions d'adaptation (figure 26 A), ainsi qu'au fil des journées pendant lesquelles se déroulent les expériences (figure 26 B).

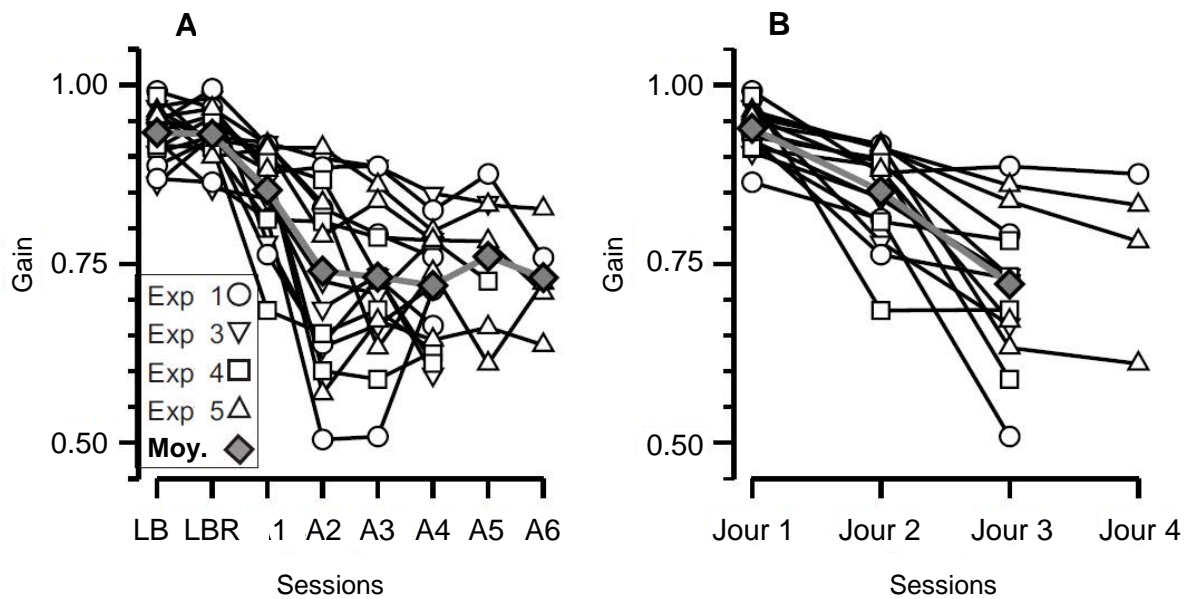


Figure 26. Rétention de la l'adaptation au cours des expériences 1 et 3 à 5. Gain saccadique moyen (calculé à partir des résultats de tous les participants, et, 'Moy.' : en moyenne) pour les 25 premiers essais : A : de chaque session et B : de chaque journée.

D'autres expériences ont mis en évidence une rétention de l'adaptation saccadique comparable. Ainsi, lorsque des singes sont placés dans l'obscurité après les sessions d'adaptation et que l'adaptation est testée le lendemain, Straube et al. (1997) et Robinson et al. (2006) observent une rétention d'une vingtaine d'heures. Chez l'être humain, Alahyane et Pélisson (2005) montrent que l'adaptation est encore significative cinq jours après la fin de l'entraînement (bien sûr sans que les individus ne restent dans l'obscurité).

Cette persistance des changements de gain peut avoir plusieurs explications. L'une pourrait concerner le fait que les saccades réalisées en laboratoire pendant les expériences d'adaptation saccadique sont des saccades très particulières.

En effet, elles sont émises vers des cibles qui apparaissent soudainement en vision périphérique, de manière prédictible. Elles sont caractérisées également par des latences (de l'ordre de 150 à 200 ms) plus courtes que celles de la majorité des saccades exécutées dans la vie quotidienne (des saccades dites d'exploration, ou « *scanning saccades* », aux latences supérieures à 250 ms), qui sont émises de manière séquentielle entre plusieurs objets stationnaires. Il est possible que les adaptations de différents types de saccades soient assez indépendantes les unes

des autres, comme le suggèrent des tests de transfert d'adaptation (Fujita et al., 2002; Hopp & Fuchs, 2010). Il reste qu'à ce jour les conclusions des études sur ce sujet sont entachées d'une absence de cohérence entre les résultats des différents laboratoires ainsi que d'une importante variabilité inter et intra-individuelle (Hopp & Fuchs, 2010).

Dans le même ordre d'idée, les saccades réalisées dans les expériences sont beaucoup plus grandes que celles qui sont exécutées dans la vie quotidienne, or les adaptations de saccades d'amplitudes différentes sont d'autant plus indépendantes que l'écart entre la taille des saccades adaptées et celle des saccades pour lesquelles le transfert est testé est important (ex : Frens & van Opstal, 1994; Semmlow et al., 1989; voir aussi Noto et al., 1999; Straube et al., 1997; pour des conclusions similaires chez le singe).

L'adaptation saccadique obtenue dans les expériences semblerait donc relativement imperméable aux effets des saccades « orthométriques » réalisées à l'extérieur du laboratoire.

Une autre explication plausible concerne la sensibilité de l'adaptation saccadique au contexte. Ainsi, il est possible d'induire simultanément et indépendamment une augmentation et une diminution du gain en fonction de contextes visuels (Herman et al., 2009), moteurs (Alahyane & Péliçon, 2004; Shelhamer & Clendaniel, 2002; Tian & Zee, 2010) ou proprioceptifs (Shelhamer & Clendaniel, 2002). Tian et Zee (2010) montrent que cette influence différentielle des indices est d'autant plus forte que les sessions d'adaptations spécifiques aux contextes sont répétées pendant trois jours consécutifs.

Même si le contexte n'a pas été manipulé explicitement dans notre expérience, le fait que les participants reviennent plusieurs fois par semaine dans la même salle, assis dans la même position en face du même écran, avec l'empreinte dentaire et la barre frontale maintenant leur tête, peut être considéré comme faisant partie d'un contexte assez bien défini pour que le comportement soit placé sous le contrôle de stimulus discriminatifs, et qu'il ne disparaisse donc pas totalement en une journée – ce qui pourrait expliquer aussi la rétention de l'adaptation observée par Alahyane et Péliçon (2005).

IV. Conclusion de la première étude : le contrôle du gain saccadique par les contingences de renforcement

Plusieurs arguments, présentés au cours de la première partie de cette thèse, remettent en cause le fait que la simple lecture de l'erreur rétinienne post-saccadique suffise à contrôler l'adaptation saccadique. Selon l'approche computationnelle, l'erreur rétinienne ferait l'objet de traitements élaborés, réalisés par des mécanismes internes spécialisés dans la calibration de gains internes (ex : Chen-Harris et al., 2008; Collins & Wallman, 2011; Ethier et al., 2008; Körding, 2007; Wong & Shelhamer, 2011).

Rappelons notamment que l'un de ces arguments repose sur l'existence d'un intervalle de temps critique au delà duquel l'apparition de la cible après son second saut n'entraîne plus de modification du gain (Bahcall & Kowler, 2000; Fujita et al., 2002). Un autre élément concerne la sensibilité de l'adaptation saccadique au contexte : un contexte particulier, qu'il soit sensoriel ou moteur, peut en effet être associé à un certain niveau de gain, et ce de manière sélective (Alahyane & Pélisson, 2004; Herman et al., 2009; Shelhamer & Clendaniel, 2002; Tian & Zee, 2010) et durable (Alahyane & Pélisson, 2005).

Nos résultats montrent que le signal qui guide l'adaptation saccadique n'est pas nécessairement de nature visuelle. Ils plaident en faveur d'une interprétation en termes de processus plus général d'apprentissage : en termes de contingences de renforcement.

Il semblerait d'ailleurs possible d'interpréter également les éléments repris ci-dessus à la lumière de l'approche opérante. Ainsi, le fait que le timing de l'erreur post-saccadique affecte l'ampleur des changements de gain saccadique pourrait être envisagé comme la nécessité de la contiguïté temporelle entre la réponse et l'agent renforçateur pour que ce dernier puisse acquérir le contrôle du comportement (Donahoe et al., 1993; Skinner, 1953). Cela permettrait d'expliquer aussi pourquoi, selon l'étude de Panouillères et al. (2011), 15 ms de présence de la cible à la fin de la saccade suffisent pour produire des changements de gain.

En outre, la sensibilité de l'adaptation saccadique au contexte pourrait traduire le contrôle du gain des saccades par des stimulus discriminatifs. Ceux-ci, en signalant la probabilité des agents renforçateurs – en l'occurrence la vision nette de la cible – influenceraient la probabilité d'émettre certaines réponses opérantes – y compris la probabilité d'émettre certaines saccades.

ETUDES 2 & 3 : RENFORCEMENT DE LA
VARIABILITE DE L'AMPLITUDE
SACCADIQUE

Si le contrôle de propriétés saccadiques tels que le gain, la latence ou encore le pic de vitesse relève d'un mécanisme général d'apprentissage (un apprentissage opérant), les conséquences des saccades devraient contrôler aussi la variabilité de ces mouvements. En effet, des travaux comme ceux de Neuringer ont mis en évidence que la variabilité d'autres comportements, tels que des séquences d'actions par exemple, pouvait dépendre des contingences de renforcement. A notre connaissance, la seule recherche consistant à renforcer différents niveaux de variabilité oculomotrice est celle de Madelain et al. (2007), qui porte sur la variabilité des latences saccadiques.

Les deux études suivantes ont pour objectif de montrer qu'une part de la variabilité de l'amplitude saccadique pourrait résulter d'un apprentissage opérant.

Dans la première des deux études, nous avons manipulé les contingences de renforcement afin d'induire expérimentalement différents niveaux de variabilité dans le gain de saccades horizontales. Cette étude fait l'objet d'un article paru dans le *Journal of the Experimental Analysis of Behavior* en 2011 ; il est placé en annexe 5.

Dans une première expérience, des agents renforçateurs auditifs ont été employés afin de tester l'étendue du contrôle qu'il est possible d'exercer sur la variabilité des gains saccadiques. Des procédures de renforcement opérant ont été appliquées pour doubler la dispersion des amplitudes, puis pour la réduire d'autant. Parallèlement, il s'agissait de maintenir la médiane des distributions à un niveau constant tout au long de l'expérience. Comme dans l'étude précédente, il était nécessaire de contrôler les conséquences visuelles des saccades, de telle sorte que l'erreur rétinienne post-saccadique soit constante. Par conséquent, le procédé de stabilisation de l'image de la cible sur la fovéa a été réutilisé.

La seconde expérience a consisté à montrer, via l'utilisation d'un groupe contrôle *yoke*, que la contingence entre les mouvements oculaires et leurs conséquences était nécessaire pour contrôler cette variabilité.

Dans la mesure où effectivement il a été possible de contrôler le niveau de variabilité du gain saccadique, il ne semble pas que l'origine de la variabilité réside uniquement dans l'existence de bruit interne incontrôlable dont les conséquences seraient à minimiser – ce que postulent notamment les théories computationnelles.

Ces résultats soulignent le rôle fonctionnel des conséquences des saccades, en l'occurrence des conséquences auditives, dans le contrôle de leur variabilité. Or une conséquence fonctionnelle se définit a posteriori, si et seulement si elle a des effets sur le comportement. La

dernière étude teste directement l'hypothèse selon laquelle la vision d'une cible peut effectivement contrôler cette propriété.

La tâche des participants n'était plus uniquement de faire des saccades horizontales, mais ils étaient placés dans une situation de recherche visuelle. Nous avons élaboré un nouveau paradigme dans lequel l'apparition de la cible était utilisée en guise d'agent renforçateur, c'est-à-dire que la cible apparaissait seulement si la variabilité des amplitudes saccadiques correspondait aux critères de renforcement. Cette étude fait l'objet de deux communications affichées (aux conférences internationales de la *Vision Sciences Society* et la *Society for Quantitative Analyses of Behavior* en mai 2012).

Dans une première expérience, il s'agissait là aussi de tester s'il était possible d'augmenter puis de baisser le niveau de variabilité en manipulant les contingences de renforcement. Enfin, la dernière expérience avait pour but de voir si le niveau de variabilité résultant de cet apprentissage opérant se transférait à des situations de recherche visuelle conventionnelle dans laquelle la cible est visible indépendamment du comportement de l'observateur.

I. Méthode de l'étude 2 : renforcement de la variabilité du gain de saccades horizontales

1.1. Participants

La première expérience a été passée par les deux auteurs (S1 et S2) ainsi qu'une étudiante de Master 2 (S3). Tous trois connaissaient l'objectif de l'étude, ils avaient déjà participé à des expériences relatives aux mouvements oculaires auparavant.

Les huit personnes (S4 à S11) qui ont participé à la seconde expérience étaient des participants naïfs qui n'avaient réalisé aucune expérience sur les mouvements oculaires avant celle-ci. Ils ont effectué une cinquantaine d'essais pour se familiariser avec le matériel et les stimulus ; les données de ces essais préliminaires ne sont pas rapportées.

Les participants, âgés de 18 à 25 ans, avaient une vision normale ou corrigée ; ils ont tous signé un formulaire de consentement (cf. annexe 1).

1.2. Equipement et stimulus

Le système d'enregistrement des mouvements oculaires (ISCAN) employé dans l'étude précédente a été utilisé dans cette étude également. De même, les caractéristiques des stimulus des deux études (croix de fixation, cible, couleur du fond, agent renforçateur sonore) étaient identiques.

1.3. Procédures des expériences

1.3.1. Expérience 1 : étendue du contrôle de la variabilité opéré par les contingences de renforcement

Afin d'évaluer l'ampleur du contrôle que pouvait exercer les contingences de renforcement sur la variabilité, quatre conditions expérimentales successives ont été mises en place : une *ligne de base ordinaire*, puis une *ligne de base stabilisée*, suivie de sessions *d'apprentissage* pour augmenter la variabilité et enfin une condition dite de *récupération* pendant laquelle les contingences de renforcement visaient à réduire le niveau de variabilité. En général, les personnes effectuaient deux sessions par jour, séparées par une pause d'une dizaine de minutes.

a. Première ligne de base

Cette expérience débutait par quatre sessions de 200 saccades qui servaient notamment à mesurer le gain saccadique initial (correspondant au gain médian de cette condition) pour chaque participant. Aucun son n'était diffusé.

Chaque essai commençait avec l'apparition de la croix de fixation, à une position variant pseudo-aléatoirement entre 2.4° et 13° à gauche ou à droite du centre de l'écran. Cette croix était présente pendant 500 à 1000 ms. Dès qu'elle disparaissait, la cible (le patch gaussien) était affichée dans la moitié opposée de l'écran ; son excentricité était comprise entre 9.5° et 14.2° . Les participants émettaient donc des saccades horizontales de droite à gauche puis de gauche à droite. Ils avaient 700 ms à partir de l'apparition de la cible pour exécuter leur saccade. Lors d'un essai, la cible restait visible pendant une durée totale de 1200 ms, puis la croix de fixation la remplaçait. Si la latence de la saccade était inférieure à 100 ms ou supérieure à 700 ms, l'essai était annulé, c'est-à-dire que la croix de fixation remplaçait immédiatement la cible, sans attendre les 1200 ms.

b. Ligne de base stabilisée

Dans cette condition expérimentale, comme dans les suivantes, l'image de la cible était stabilisée sur la rétine de telle sorte que la cible s'y projette à une certaine distance de la fovéa. Cette distance restait constante tout au long de l'expérience et dépendait du gain saccadique médian de chaque individu mesuré lors de la condition précédente. Ainsi, dès que la saccade était détectée, la cible disparaissait puis elle était affichée à la fin du mouvement à une nouvelle position obtenue en multipliant le point d'arrivée de l'œil (par exemple 12°) par l'inverse du gain initial (par exemple 0.9°). Par conséquent, si un participant avait un gain initial hypométrique, la cible continuait à apparaître légèrement plus loin de la position de l'œil à la fin de la saccade, de manière à reproduire cette hypométrie.

Dans l'exemple donné, la cible serait donc affichée à $12 \times (1/0.9) = 13.33^\circ$ à droite de la croix de fixation initiale, comme le représente la double flèche sur la figure 27 A. Cette manipulation avait pour objectif d'éviter la diminution du gain observée dans l'étude précédente – ainsi que dans d'autres expériences utilisant ce procédé de stabilisation de l'image de la cible, comme celles de Havermann et Lappe (2010), Robinson et al. (2003) ou Wong et Shelhamer (2011). Puis, un nouvel essai commençait avec la réapparition de la croix de fixation qui venait remplacer la cible.

La médiane du gain obtenue sur ces 800 saccades « stabilisées » était alors utilisée comme médiane cible dans les conditions expérimentales suivantes. C'est cette ligne de base qui a été prise comme référence dans les analyses des résultats.

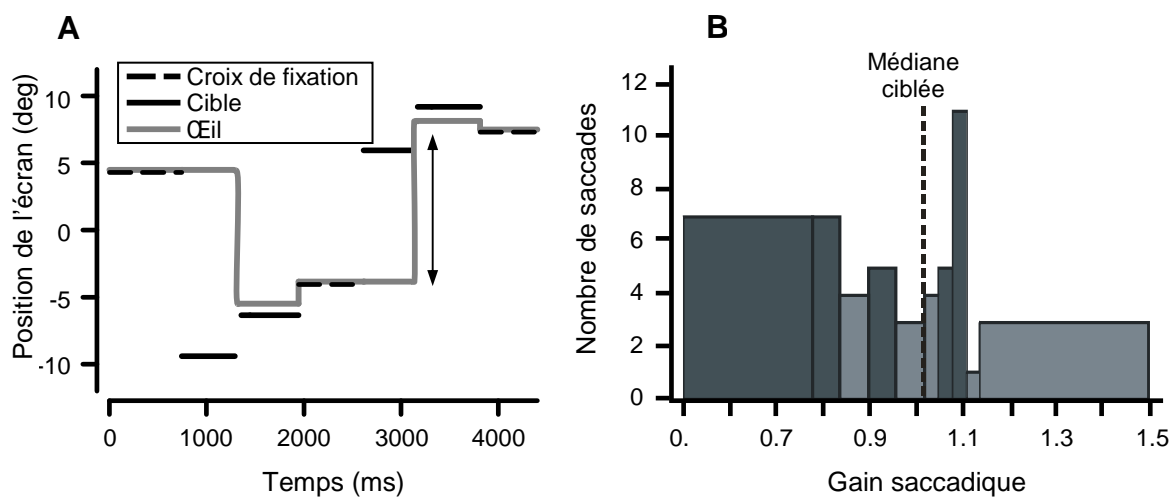


Figure 27. A : exemples représentatifs des positions de la croix de fixation, de la cible et de l'œil ; la double flèche indique la saccade prise en exemple dans le texte. B : classes des gains et critères de renforcement pour un essai réalisé par un participant (S2). Un agent renforçateur auditif (un son bref) était émis si le gain de la saccade réalisée tombait dans l'une des cinq classes les moins fréquentes (gris clair).

c. Apprentissage

L'augmentation de la variabilité des amplitudes saccadiques a été induite grâce à une procédure de renforcement différentiel des amplitudes les moins fréquentes (les amplitudes les plus fréquentes étant soumises à extinction), une procédure adaptée de celles de Blough (1966) et de Denney et Neuringer (1998).

Pour déterminer les critères de renforcement propres à chaque participant, 10 classes de gains étaient calculées, centrées sur le gain de la ligne de base stabilisée. Ces classes étaient inégales à cause de la tendance à l'hypométrie généralement observée en laboratoire (Becker, 1989; Henson, 1978; Irving et al., 2006) : les cinq classes inférieures avaient une amplitude de 0.06 unité de gain tandis que celle des cinq classes supérieures était de 0.03 unité de gain. Les valeurs extrêmes de l'histogramme étaient fixées respectivement à des gains de 0.5 et 1.5. Par exemple, comme l'illustre la figure 27 B, pour le participant S2 dont la médiane du gain mesurée en ligne de base était 1.02, les intervalles des différentes classes étaient [0.5 ; 0.78[; [0.78 ; 0.84[etc., les dernières classes étant délimitées par les intervalles [1.08 ; 1.11[; [1.11 ; 1.5].

A chaque essai, la fréquence des différentes classes d'amplitudes était déterminée à partir des 50 saccades précédentes (chaque nouvelle saccade remplaçant la plus ancienne). Pour les premiers essais d'une session, les fréquences étaient obtenues à partir de 50 essais tirés aléatoirement dans la session précédente. Si le gain de la saccade actuelle était compris dans l'une des cinq classes dont la fréquence était la plus basse, un son était émis.

De plus, comme dans l'étude précédente, dès que trois essais sur cinq étaient suivis d'un son, un autre agent renforçateur auditif était délivré : une voix synthétique prononçait « bravo », indiquant que le participant avait gagné un point.

Le nombre de sessions (de 200 essais) d'apprentissage dépendait du temps mis pour doubler l'écart-type de la distribution des gains observés en ligne de base stabilisée. Les participants S1, S2 et S3 ont respectivement effectué 40, 18 et 32 sessions, à raison de quatre à six sessions par jour, séparées d'une dizaine de minutes.

d. Retour au niveau de base (ou récupération)

Lors de cette quatrième condition expérimentale (nommée *récupération* par souci de concision, traduction du terme anglais de « *recovery* »), il s'agissait de réduire le niveau de variabilité.

Un programme de renforcement percentile ciblé a été appliqué de la même manière que dans l'étude précédente, à l'exception de deux éléments. Le premier concernait évidemment la détermination de l'amplitude cible : celle-ci correspondait au gain médian mesuré en ligne de base stabilisée. Le second était le pourcentage de réponses renforcées par le programme, qui était de 30% dans cette étude.

Le nombre de sessions nécessaires pour réduire le niveau de variabilité a été respectivement de 18, 22 et 6 pour les participants S1, S2 et S3.

1.3.2. Expérience 2 : la contingence entre les mouvements oculaires et leurs conséquences est-elle nécessaire pour contrôler la variabilité ? *Yoke control*

L'utilisation d'un groupe contrôle *yoke* était capitale pour achever de montrer que ni l'intermittence du renforcement (ex. : Lee et al., 2007; Machado, 1989; Neuringer, 2002, 2009), ni les instructions (Hopkinson & Neuringer, 2003; Souza et al., 2012) ni le procédé de stabilisation de l'image de la cible n'étaient responsables des changements de variabilité induits par les procédures comportementales.

Les huit participants ont reçu comme consigne de faire une saccade vers la cible dès que la croix de fixation disparaissait ainsi que d'entendre le plus de sons possible, afin de gagner le maximum de points et de remporter le lot du vainqueur. A l'issue des deux lignes de base, il leur était précisé que les sons qu'ils entendaient dépendaient de la variabilité de l'amplitude de leurs saccades (la totalité des consignes est retranscrite en annexe 3). Néanmoins, aucune explication n'était donnée relativement à notre intention d'augmenter la variabilité ni à la stabilisation de l'image de la cible.

Seules les quelques modifications qui ont été apportées à la procédure de l'expérience précédente sont décrites ci-dessous. Tous les autres éléments étaient identiques.

a. Première ligne de base

Au début de chaque essai, la croix de fixation réapparaissait à une distance aléatoire comprise entre $\pm 2.4^\circ$ et $\pm 13^\circ$ du centre de l'écran (du même côté que la cible), et non à la place de cette cible. Ceci était valable pour les conditions suivantes.

b. Ligne de base stabilisée

Pour chacun des huit participants, le programme de renforcement percentile était mis en place lors de cette seconde ligne de base. Ainsi les 30% des saccades les plus proches de la valeur initiale (soit la médiane obtenue à partir des saccades de la première ligne de base) étaient renforcées par un son. Comme précédemment, la médiane du gain de cette seconde ligne de base servait ensuite de valeur centrale pour le calcul des critères de renforcement.

c. « Apprentissage »

Les participants étaient répartis en deux groupes. Pour ceux du groupe expérimental (S4 à S7), la procédure de renforcement des amplitudes les moins fréquentes décrite ci-dessus était appliquée. Elle visait à multiplier par 1.5 l'écart-type de la distribution obtenue lors de la ligne de base stabilisée.

Quant aux participants du groupe contrôle (S8 à S11), les conséquences auditives de leurs saccades étaient indépendantes de l'amplitude de leurs mouvements : elles étaient en effet programmées à l'avance en fonction de ce qu'avait obtenu l'un des participants du groupe contrôle. Par exemple, si le participant du groupe expérimental S4 avait entendu un son à la fin de ses deuxième et quatrième essais, le participant contrôle correspondant entendait lui aussi un son à l'issue de ses deuxième et quatrième essais, quelle que soit l'amplitude de ses saccades. Ainsi, les positions des croix de fixation et des cibles ainsi que les séquences de renforcement présentées à chaque participant du groupe expérimental étaient répliquées pour l'un des participants du groupe contrôle. Le procédé de stabilisation de l'image de la cible dépendait toutefois du niveau de gain des participants contrôles, et les essais dont la latence était soit inférieure à 100 ms soit supérieure à 700 ms étaient également annulés. Même si le terme « d'apprentissage » peut paraître inapproprié pour ces participants, il sera conservé pour faire référence à cette condition expérimentale.

Si la contingence entre les conséquences des mouvements oculaires et la variabilité est nécessaire pour induire des changements de variabilité, alors ces changements ne devraient correspondre à ce qui était attendu que pour les participants du groupe expérimental.

Les participants S4 à S7, ainsi que leurs participants *yoke* respectifs, ont réalisé 17, 18, 10 et 22 sessions (de 200 essais) lors de cette condition.

1.4. Acquisition et analyse des données

L'acquisition et l'analyse des données *online* n'étaient pas très différentes de celles réalisées dans l'étude 1. Seuls les quelques détails qui ont divergé sont donc présentés ici. En premier lieu, la fenêtre de détection et de mesure *online* des saccades était de 700 ms, comptabilisées à partir de l'apparition de la cible. En second lieu, les saccades dont le gain était soit inférieur à 0.5, soit supérieur à 1.5 ou encore celles dont la latence était inférieure à 100 ms étaient écartées des analyses *offline* – en moyenne, 3,27% des saccades ont été éliminées.

La médiane et l'écart-type des gains saccadiques ont été déterminées sur les 800 dernières saccades des différentes conditions expérimentales. Un second indice de variabilité a aussi été calculé : la *U value* (ou « *uncertainty* »), qui renseigne sur l'incertitude d'une distribution. La formule et la signification de cette mesure, couramment utilisée pour rendre compte de la variabilité comportementale (ex. : Machado, 1989; Neuringer, 2002; Page & Neuringer, 1985), sont explicitées à la fin de la partie dédiée aux résultats de cette étude.

Enfin, le pic de vitesse et la latence des saccades ont également fait l'objet d'analyses.

Afin de comparer individuellement les différentes conditions, les intervalles de confiance percentiles (95%) ont été calculés (sur les 800 dernières saccades de chaque condition expérimentale) avec la méthode du Bootstrap (Efron, Jolivet, & Hordan, 1995) pour chaque participant. Celle-ci présente l'avantage de ne reposer sur aucune hypothèse relative à la distribution.

II. Résultats de l'étude 2

Les distributions des gains saccadiques ont été modifiées en fonction des contingences de renforcement. Pour chacun des participants des deux groupes expérimentaux, une augmentation de la dispersion du gain, mais pas de la médiane, a été systématiquement observée lors de l'apprentissage. Quant au groupe contrôle, les changements éventuels de leurs distributions ne suivaient aucune tendance précise.

2.1. Expérience 1 : étendue du contrôle de la variabilité opéré par les contingences de renforcement

Sur la figure 28 A est indiqué, pour le participant S2, le gain de chaque essai lors de chaque condition expérimentale. La figure 28 B résume ces données : les distributions des fréquences des gains saccadiques y sont représentées par des histogrammes obtenus à partir des quatre dernières sessions (soit 800 essais) de chaque condition expérimentale.

A la fin de l'apprentissage, la dispersion des gains saccadiques a nettement augmenté par rapport à la ligne de base. Puis elle a été réduite à la fin de la récupération : les distributions de la ligne de base et de la récupération étaient quasiment superposables. L'écart-type est passé de 0.08 lors de la ligne de base stabilisée à 0.2 à la fin de l'apprentissage et il est revenu à 0.08 lors des dernières sessions de récupération. Quant à la médiane des gains, elle est restée globalement inchangée pendant les différentes conditions expérimentales (où elle était respectivement de 1.02, 0.98 et 1.04).

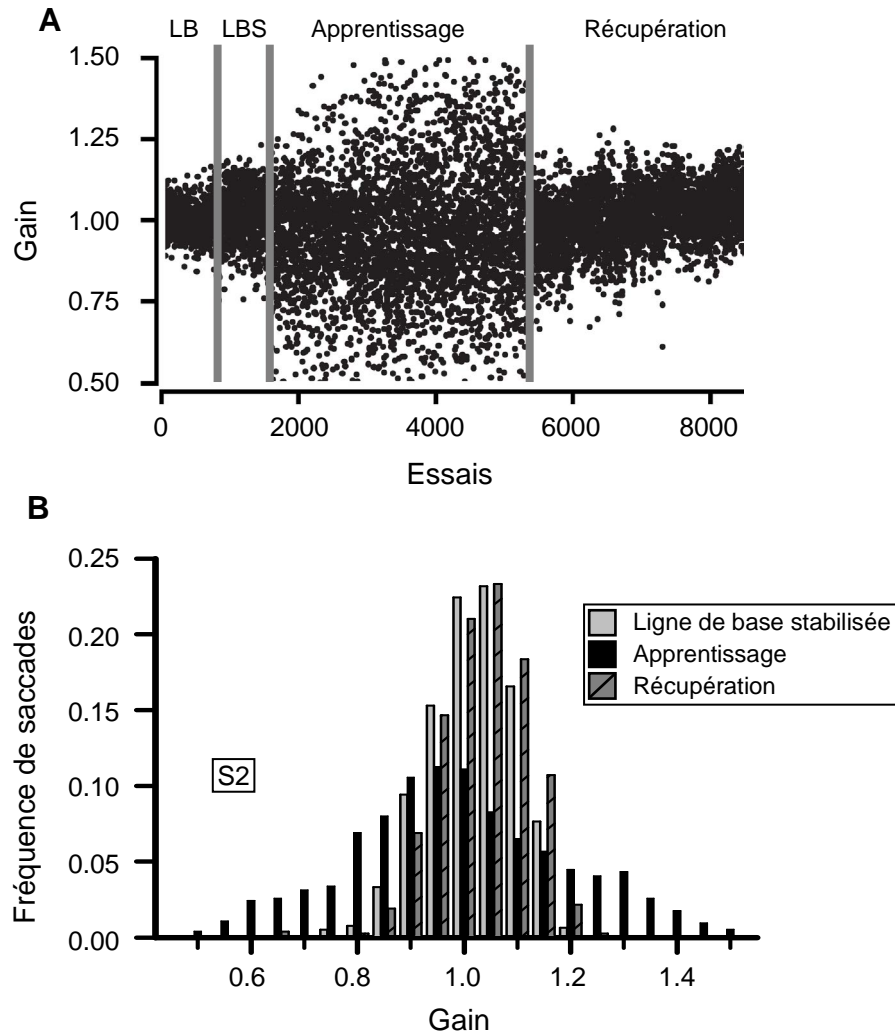


Figure 28. Résultats de l'expérience 1 pour le participant S2. A : gain saccadique de chaque essai dans les différentes conditions expérimentales, un point correspondant à un essai (LB : ligne de base, LBS : ligne de base stabilisée). B : distribution des fréquences des 800 derniers gains saccadiques de chaque condition.

La figure 29 présente les résultats (avec les intervalles de confiance Bootstrap) des trois participants. Comme pour S2, à la fin de l'apprentissage l'écart-type du gain des deux autres participants était significativement supérieur à ceux des deux autres conditions (figure 29 A). En moyenne, entre la ligne de base et les quatre dernières sessions d'apprentissage, l'écart-type du gain a été multiplié par 2.21. En revanche, entre la ligne de base et la récupération, il a été multiplié par 1.06 : le niveau de variabilité est donc presque revenu à son niveau de base – les différences éventuelles entre ces deux conditions ne faisant apparaître aucune tendance précise.

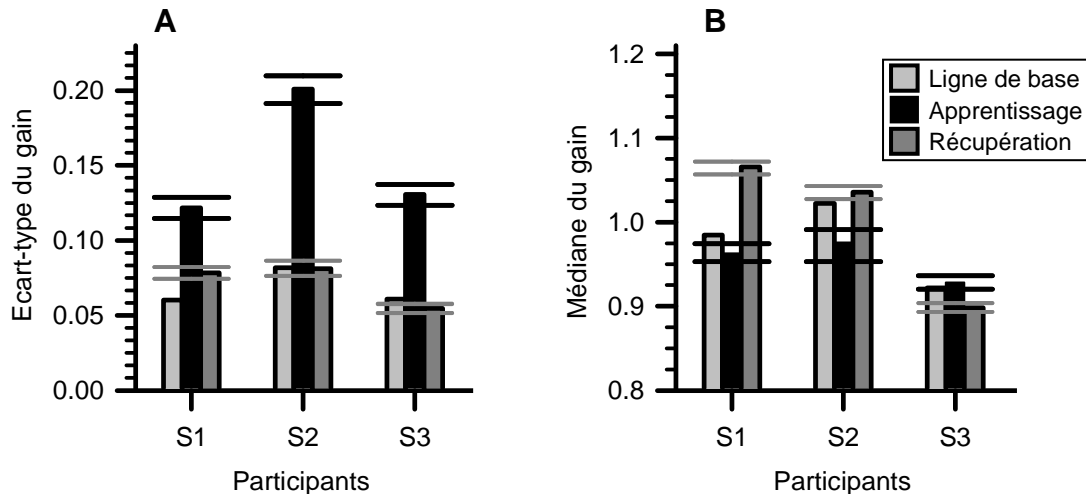


Figure 29. Résultats de l'expérience 1 pour les trois participants. A : écart-type des gains saccadiques pour les quatre dernières sessions (800 essais) de chaque condition expérimentale. Les intervalles de confiance Bootstrap (95^{ème} percentile) sont indiqués par les traits horizontaux, noirs pour l'apprentissage et gris pour la récupération. B : médiane et intervalle de confiance des gains saccadiques correspondants.

Les médianes des gains ont été un peu altérées, mais de manière inégale selon les participants (figure 29 B). Pour S1 et S2, la médiane la plus faible était celle des dernières sessions d'apprentissage tandis que pour S3, la médiane de la ligne de base et de l'apprentissage étaient égales, elle a baissé uniquement lors de la récupération. En moyenne, entre la ligne de base et la fin de l'apprentissage, la médiane a été multipliée par 0.98, et par 1.02 entre l'apprentissage et la récupération.

2.2. Expérience 2 : la contingence entre les mouvements oculaires et leurs conséquences est-elle nécessaire pour contrôler la variabilité ? *Yoke control*

Pour les participants du groupe expérimental, les changements des distributions des gains saccadiques étaient analogues à ceux obtenus pour les participants de l'expérience précédente. Comme l'illustre la figure 30 A, pour le participant S7, la dispersion a augmenté après

l'apprentissage (l'écart-type est passé de 0.08 à 0.14) tandis que la médiane du gain est restée identique (elle était de 0.94 lors de la ligne de base comme à la fin de l'apprentissage).

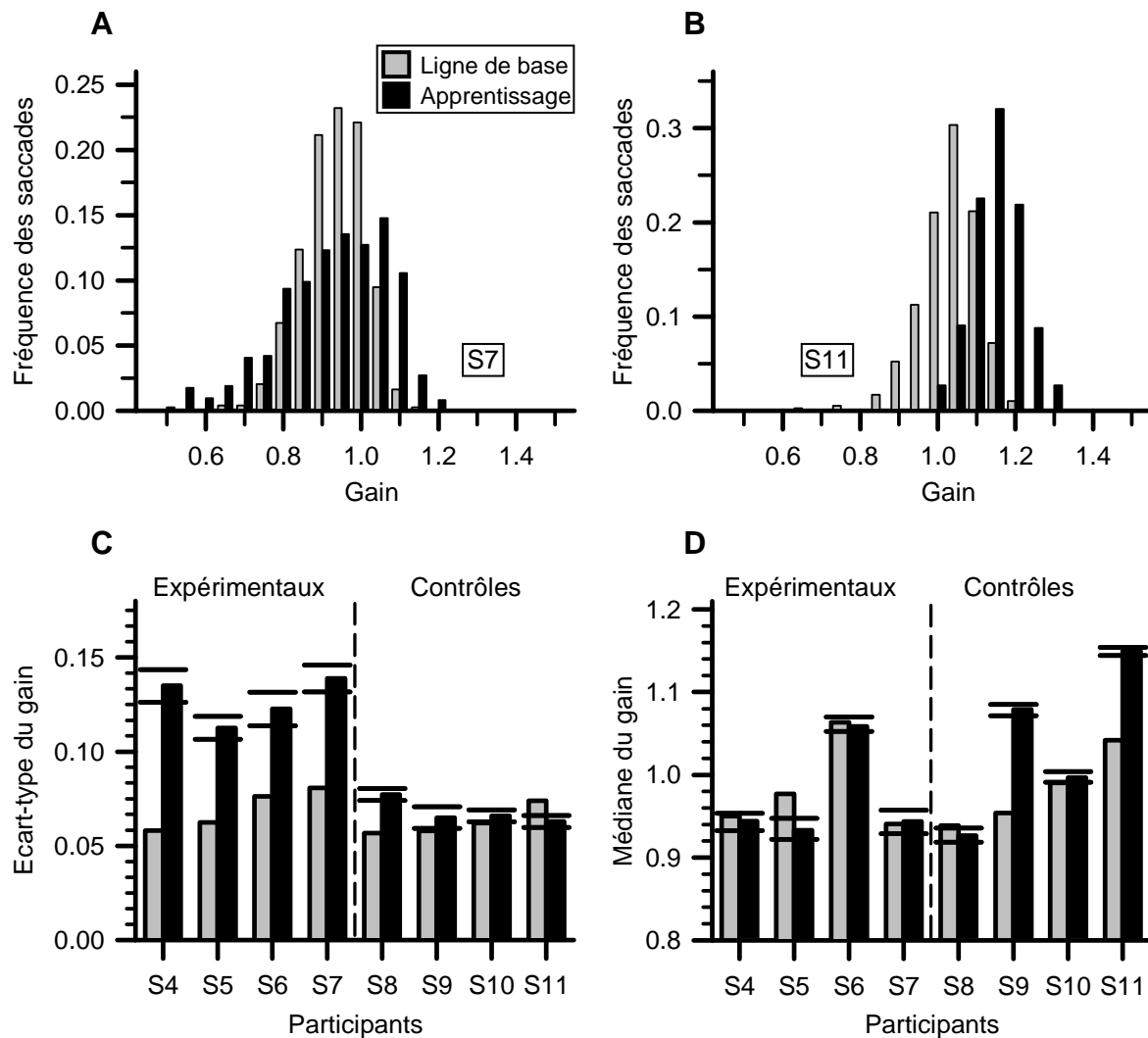


Figure 30. Résultats de l'expérience 2. A : distributions des fréquences des gains saccadiques de la ligne de base et des quatre dernières sessions (800 essais) de l'apprentissage, pour le participant S7 du groupe expérimental. B : distributions des fréquences des gains de la ligne de base et des quatre dernières sessions de la condition de renforcement non contingent (barres noires) pour le participant contrôle correspondant, S11. C : écart-type et intervalle de confiance Bootstrap (95^{ème} percentile) des gains saccadiques de chaque condition pour les participants des deux groupes, expérimental et contrôle. D : médiane et intervalle de confiance de ces gains saccadiques.

Ce n'était pas le cas pour le participant contrôle correspondant (S11, figure 30 B). L'écart-type de la distribution de ses gains saccadiques a légèrement diminué entre la ligne de base et la fin de la seconde condition expérimentale où le renforcement était indépendant de l'amplitude de ses mouvements oculaires – il est passé de 0.07 à 0.06. Quant à la médiane, de 1.04 en ligne de base, elle a atteint 1.15 lors des quatre dernières sessions de l'expérience.

Les intervalles de confiance présentés sur la figure 30 C révèlent que pour tous les participants du groupe expérimental, les écarts-types étaient significativement plus élevés à la fin de l'apprentissage que lors de la ligne de base (ils ont été multipliés par 1.86 en moyenne). En ce qui concerne les participants du groupe contrôle, si l'écart-type a été modifié par la condition de renforcement non contingent, ces changements n'étaient toutefois pas cohérents. En effet, il a augmenté pour les participants S8, S9 et S10 mais a diminué pour le participant S11. De plus, ces changements étaient beaucoup plus faibles que ceux du groupe expérimental : le facteur d'augmentation entre la ligne de base et la condition de renforcement non contingent était de 1.1.

En ce qui concerne les médianes du gain, l'histogramme de la figure 30 D montre qu'elle est restée constante pour trois participants sur quatre dans le groupe expérimental, comme l'indiquent les intervalles de confiance. La médiane n'a significativement diminué (de 0.98 à 0.93) que pour un seul participant, S5. Quant au groupe contrôle, elle a augmenté de manière significative pour les participants S9, S10 et S11, tandis qu'elle a diminué pour S8. De surcroît, les hausses observées pour S9 et S11 étaient très importantes (plus de 10%).

2.3. Analyse des *U values*

La *U value*, mesure de l'incertitude d'une distribution souvent utilisée dans les études sur la variabilité comportementale, est calculée selon la formule suivante :

$$U = \frac{-\sum_{n=1}^N (p_n \log_2(p_n))}{\log_2(N)} \quad (5)$$

dans laquelle p représente la fréquence relative d'une classe n et N le nombre total de classes. Dans ces deux expériences, pour calculer la *U value*, 10 classes d'amplitude équivalente ont été utilisées (des classes de 0.1 unité de gain, allant de 0.5 à 1.5). La *U value* reflète la vraisemblance que le gain d'une saccade tombe dans chacune des classes. Si les 10 classes contiennent toutes un nombre équivalent d'observations, alors $U = 1$. En revanche, si toutes les observations sont regroupées dans une seule des classes, $U = 0$.

La *U Value*, qui ne requiert aucune hypothèse quant à la forme de la distribution, s'avère un bon complément à l'écart-type pour décrire la variabilité des distributions finies des classes d'amplitudes, comme l'enseigne la figure 31. Contrairement à l'écart-type, la *U Value* n'est pas sensible à la valeur de la modalité des classes (figures 31 A et B), mais elle est avant tout sensible à la fréquence relative de la classe. La comparaison des figures 31 A et C illustre le fait que l'écart-type peut diminuer alors que le comportement est plus variable, ce qui n'est pas le cas de la *U Value*. Celle-ci est en effet affectée par la proportion de classes vides (figures 31 C, D, E et F). Cette proportion étant corrélée au nombre de valeurs considérées, celui-ci exerce donc aussi une influence sur cette mesure.

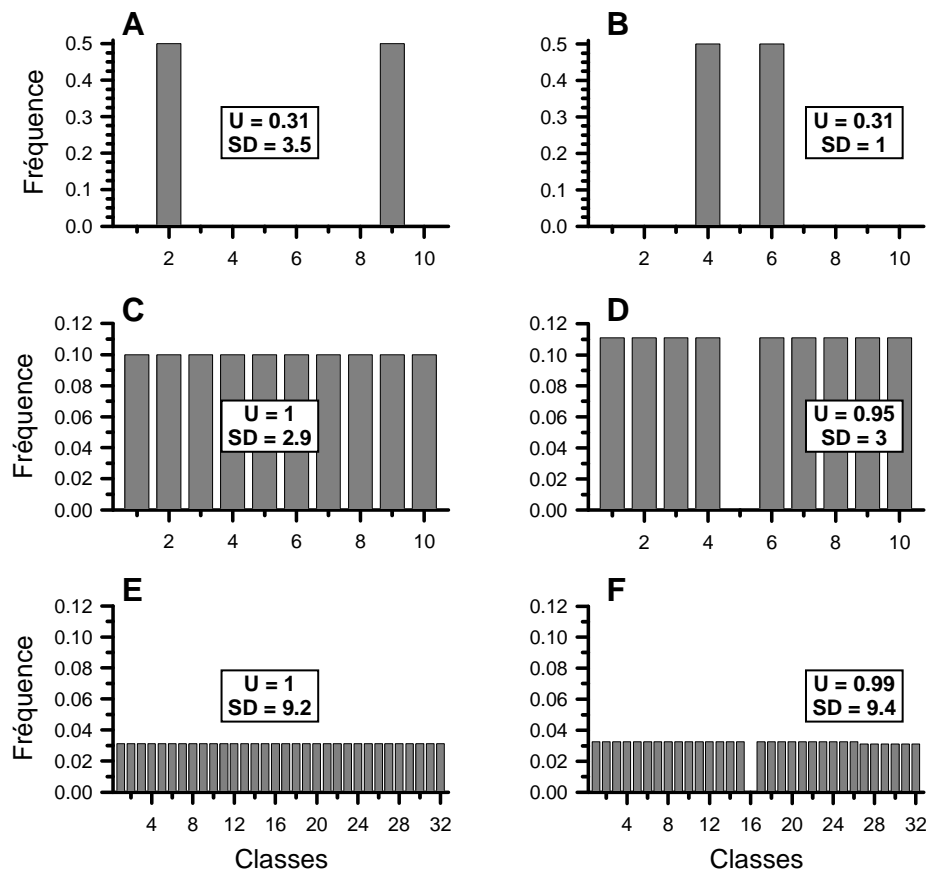


Figure 31. Comparaison de l'écart-type et de la *U Value* pour plusieurs distributions fictives, catégorisées en 10 ou 32 classes.

La figure 32 (carrés et cercles) trace les *U values* et leur intervalle de confiance Bootstrap calculés sur les quatre dernières sessions de renforcement en fonction des valeurs de la ligne de

base dans les deux expériences. Quelle que soit la condition expérimentale, les valeurs de la ligne de base étaient proches les unes des autres (comprises entre 0.41 et 0.55).

Deux ensembles de points se distinguent nettement sur le graphique. D'une part les données des participants des groupes expérimentaux (carrés et cercles blancs) apparaissent nettement au dessus de la ligne correspondant à l'absence de changement entre les deux conditions, ce qui indique une augmentation importante de la *U Value* à la fin de l'apprentissage. En effet, les valeurs ont été multipliées par des facteurs moyens de 1.61 (expérience 1) et 1.49 (pour le groupe expérimental de l'expérience 2).

D'autre part, les données du groupe contrôle (symboles gris) sont proches voire en dessous de la ligne : peu de changements ont eu lieu entre la ligne de base et la condition de renforcement non contingent. En moyenne, la *U Value* a été multipliée par 1.08 entre ces deux conditions, mais elle est restée constante pour l'un des participants (S10) et a diminué pour un autre (S11). Quant aux données de la récupération (triangles blancs, expérience 1), elles sont très proches de celles des participants contrôles : entre la ligne de base et la récupération, elles n'ont été multipliées que par 1.05 en moyenne. Pour deux participants sur trois, la *U Value* était équivalente voire plus faible à la fin de la récupération par rapport à la ligne de base (ce qui est confirmé par les intervalles de confiance).

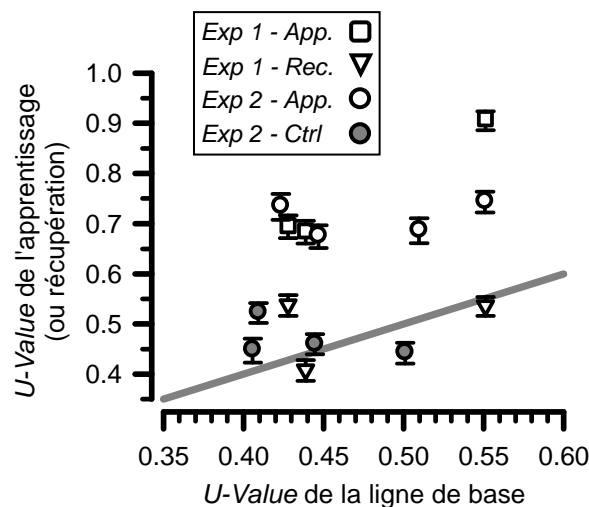


Figure 32. *U Value* moyenne lors de la ligne de base en fonction de la *U Value* moyenne calculée sur les quatre dernières sessions de l'apprentissage des groupes expérimentaux (carrés et cercles blancs) ou de la récupération (triangles blancs). Cercles : données de l'expérience 2. Symboles gris : données des participants du groupe contrôle. Les barres représentent les intervalles de confiance Bootstrap (95^{ème} percentile) de la *U Value* de l'apprentissage.

III. Discussion de l'étude 2

Cette étude montre qu'une part de la variabilité de l'amplitude saccadique peut résulter d'un apprentissage opérant.

La première expérience atteste que le renforcement peut induire des changements importants dans la dispersion des distributions de gains saccadiques tout en maintenant la médiane à un niveau constant. Ainsi, par rapport à la ligne de base stabilisée, l'écart-type est multiplié par 2.21 en moyenne à la fin de l'apprentissage. De plus, suite à la modification des contingences de renforcement, la variabilité redevient similaire à celle du niveau de base. Par conséquent, ces effets ne peuvent pas être attribués au procédé de stabilisation de l'image de la cible sur la fovéa.

Les instructions (ou, pour les auteurs, la connaissance des règles de programmation) auraient pu jouer également un rôle non négligeable dans les programmes de renforcement de la variabilité, comme l'ont évoqué précédemment plusieurs auteurs (Hopkinson & Neuringer, 2003; Lattal & Neef, 1996; Souza et al., 2012). Cependant, dans la seconde expérience de cette étude, les données des participants du groupe contrôle indiquent que les instructions ne peuvent expliquer à elles seules l'augmentation systématique de la variabilité obtenue dans les groupes expérimentaux. En effet, de tels changements ne s'observent pas quand le renforcement n'est pas contingent à cette dimension.

Ces expériences prolongent les résultats de l'étude de Madelain et al. (2007) qui montrent que la variabilité des latences saccadiques peut être contrôlée directement par les contingences de renforcement et qu'elle peut être manipulée indépendamment de la médiane de ces latences. Par ailleurs, elles contrastent avec les travaux de l'approche computationnelle classique qui affirme que la variabilité sensori-motrice provient exclusivement d'un bruit interne stochastique et incontrôlable (Faisal et al., 2008; van Beers, 2007).

Si les procédures de renforcement appliquées dans cette étude ont pu modifier la variabilité *opérante* du gain saccadique, n'altèreraient-elles pas parallèlement les autres propriétés des saccades ? Cette discussion examine d'abord l'influence éventuelle que la fréquence de renforcement pourrait exercer indirectement sur la variabilité. Puis elle aborde l'analyse des latences d'une part, et celle de la *main sequence* d'autre part.

3.1. L'influence éventuelle de la fréquence de renforcement

Des études réalisées sur des séquences d'actions chez des animaux soulignent que la fréquence de renforcement dans les programmes utilisés peut avoir un effet sur la variabilité comportementale, et donc interagir avec les programmes de renforcement opérant, contingent à cette variabilité (Lee et al., 2007; Neuringer, 2002, 2009). Même si cette influence semble moins importante que celle du renforcement direct, opérant (Grunow et Neuringer, 2002), elle justifie l'emploi de procédures telles que le renforcement percentile qui permet de dissocier ces deux effets (Machado, 1989, 1992).

A l'instar de Page et Neuringer (1985), nous avons utilisé un groupe contrôle *yoke* afin de montrer que la fréquence de renforcement ne suffit pas pour rendre compte des changements de variabilité constatés. Cependant, cela n'équivaut pas au contrôle opéré par une procédure de renforcement percentile, et c'est ce qui pourrait expliquer notamment les différences inter-individuelles observées dans les groupes expérimentaux. En effet, lors de l'expérience 1, les *U Values* vont de 0.69 à 0.91 ; dans l'expérience 2, elles vont de 0.68 à 0.75, alors que le programme de renforcement opérant est identique. De surcroît, le nombre de sessions nécessaires pour multiplier l'écart-type respectivement par 2 ou 1.5 était lui aussi différent en fonction des personnes, autre signe possible d'effets non maîtrisés du taux de renforcement.

Pour voir dans quelle mesure la fréquence de renforcement pourrait influencer la variabilité, les coefficients de corrélation entre la *U value* et le taux de renforcement des différentes sessions ont été déterminés pour chaque participant. Ils sont indiqués dans le tableau 1.

Les trois premières lignes correspondent aux données de l'expérience 1. Pour deux participants sur trois, les corrélations sont positives, ce qui est attendu, puisque le renforcement est contingent à la variabilité – mais regrettable dans une expérience portant sur la variabilité opérante, qui devrait donc faire en sorte de maintenir le taux de renforcement constant.

Que dire du participant S2, pour qui la corrélation entre la *U value* et le taux de renforcement est négative ? Cela indiquerait-il que, conformément aux données de la littérature (Antonitis, 1951; Cherot et al., 1996; Gharib et al., 2004; Neuringer et al., 2001; Stahlman et al., 2010), la raréfaction du renforcement augmenterait la variabilité de ce participant en particulier, et augmenterait ainsi les effets du renforcement opérant, direct, de la variabilité ? En effet, la *U Value* de ce participant à la fin de l'apprentissage est la plus élevée de toutes.

Tableau 1. Coefficients de corrélation entre la U value et le taux de renforcement des différentes sessions, calculés à partir des données de l'apprentissage pour chaque participant, et sur l'ensemble de l'expérience pour les participants de l'expérience 1. En gris : corrélation correspondant à l'hypothèse selon laquelle moins le renforcement est fréquent, plus la variabilité est élevée.

	Participant	r apprentissage	p	r apprentissage + récupération	p
Exp. 1 Groupe expérimental	S1	0.839	< 0.001	0.754	0.0000
	S2	-0.725	< 0.001	0.008	0.9581
	S3	0.471	0.007	0.449	0.005
Exp. 2 Groupe expérimental	S4	0.442	0.076	-	-
	S5	0.280	0.260	-	-
	S6	0.399	0.254	-	-
	S7	0.681	< 0.001	-	-
Exp. 2 Groupe contrôle	S8	0.293	0.253	-	-
	S9	0.074	0.772	-	-
	S10	-0.447	0.195	-	-
	S11	0.131	0.561	-	-

La figure 33 A présente les évolutions de la variabilité et du taux de renforcement de ce participant au fil des sessions. Pendant la récupération, lorsque le programme de renforcement percentile est appliqué, cet effet n'est plus retrouvé (le taux de renforcement – de 0.31 en moyenne – est alors constant et il n'est pas plus élevé que lors des sessions d'apprentissage – où il est de 0.35) si bien que la corrélation calculée sur la totalité des sessions n'est pas significative. Dès lors, il semble difficile d'attribuer à l'intermittence du renforcement dans cette expérience un effet sur la variabilité comportementale, et cela indique que la variabilité dépend avant tout des contingences de renforcement.

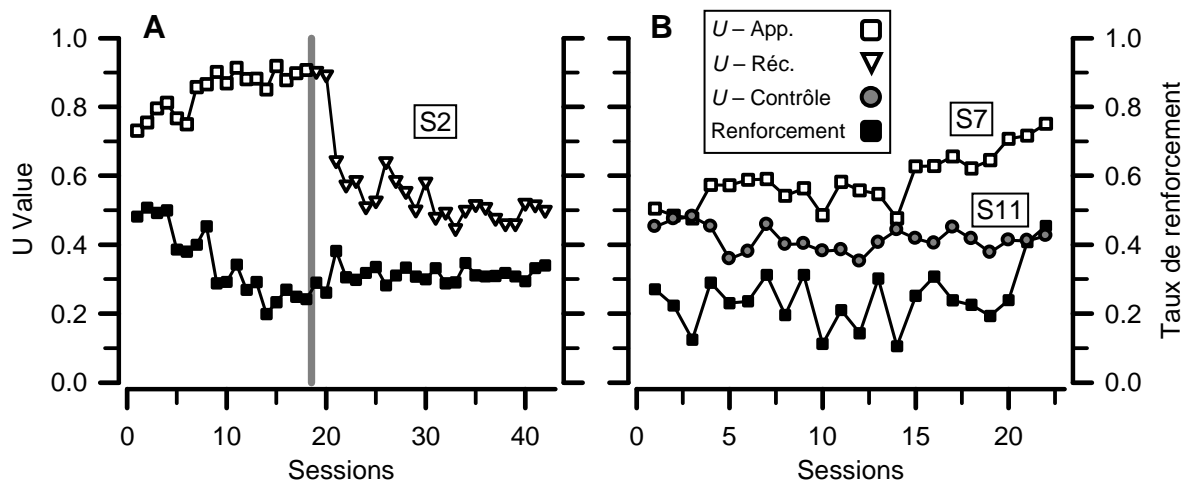


Figure 33. Evolutions de la *U Value* et du taux de renforcement au fil des sessions : A : pour le participant S2 de l'expérience 1, le seul pour qui la corrélation entre les deux variables lors de l'apprentissage est négative. La ligne verticale matérialise le changement de condition expérimentale. B : pour le participant S7, pour qui cette corrélation est positive, et le participant contrôle S11 correspondant.

En ce qui concerne l'expérience 2, le tableau 1 indique que les corrélations entre la *U value* et le taux de renforcement ne sont pas significatives. Cette différence avec l'expérience précédente pourrait s'expliquer par le fait que dans l'expérience 2, les participants sont naïfs. Seul le coefficient de corrélation du participant S7 est positif, reflétant ici aussi la contingence de renforcement.

Les données de ce participant S7, ainsi que le participant contrôle correspondant (S11), sont représentées sur la figure 33 B. Elles montrent que le taux de renforcement n'est – malheureusement – pas constant, mais que globalement il n'augmente pas (excepté à la fin de l'expérience). Ces observations sont valables aussi pour les autres participants de cette expérience. En revanche, l'évolution de la *U Value* est plus régulière, ce qui témoigne que la variabilité dépend davantage des contingences de renforcement que de la fréquence de ce dernier. Un autre argument est fourni par les données du participant contrôle, pour qui les fluctuations du taux de renforcement ne se reflètent pas sur la *U Value*.

Même si elles ne permettent pas d'expliquer les différences inter-individuelles ni de déceler une influence du taux de renforcement sur la *U Value*, ces analyses confirment que les changements de variabilité sont bien attribuables aux contingences de renforcement opérant.

3.2. Analyse des latences saccadiques

Comme dans l'étude précédente, l'examen des temps de réaction pourrait éventuellement permettre de faire quelques inférences à propos de l'influence éventuelle de stratégies cognitives (Glimcher, 2003; Reddi & Carpenter, 2000; Stritzke et al., 2009). Afin de tester les effets de la procédure de renforcement sur les latences saccadiques pour chaque participant, les proportions de changement des latences et des *U Values* entre la dernière session du renforcement et la dernière session de ligne de base stabilisée ont été calculées :

$$\frac{\text{latence médiane (ou } U \text{) du renforcement - latence médiane (ou } U \text{) de ligne de base}}{\text{latence médiane (ou } U \text{) de ligne de base}} \times 100 \quad (6)$$

La figure 34 trace ces proportions de changement en fonction de celles des *U Values* et révèle que les latences sont globalement plus élevées après le renforcement (elles passent de 179 à 277 ms en moyenne, test de Wilcoxon bilatéral, $p = 0.005$).

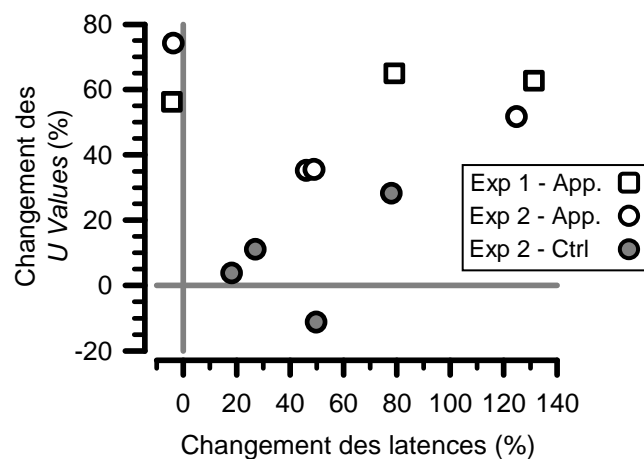


Figure 34. Proportion de changement des *U Values* en fonction de la proportion de changement des latences saccadiques (proportions moyennes calculées à partir des saccades de la dernière session de renforcement par rapport à celles de la dernière session de ligne de base stabilisée). Symboles vides : participants expérimentaux, symboles gris : participants du groupe contrôle.

Néanmoins, pour les participants des groupes expérimentaux, le coefficient de corrélation entre ces deux mesures est de -0.057 ($p = 0.903$), ce qui indique que les changements de variabilité des amplitudes saccadiques ne sont pas corrélés aux changements des latences. De plus, les groupes expérimentaux et contrôle diffèrent quant à leurs taux de changement des *U Values* mais pas pour ceux de leurs latences. Par conséquent, même si l'utilisation éventuelle de stratégies ne peut être écartée, aucune ne peut être spécifiquement associée à l'augmentation de la variabilité de l'amplitude saccadique.

3.3. Analyse de la *main sequence*

Etant donné qu'un certain nombre d'études, y compris l'étude précédente sur le façonnage du gain saccadique, suggèrent que le renforcement différentiel des saccades peut altérer la *main sequence* des mouvements renforcés chez l'être humain (Madelain & Paeye, 2011; Montagnini & Chelazzi, 2005; Xu-Wilson et al., 2009), la procédure de renforcement de la variabilité des saccades pourrait elle aussi avoir une influence sur cette propriété.

De la même manière que dans l'étude précédente, nous avons donc comparé la *main sequence* des saccades enregistrées lors de la première ligne de base (non stabilisée) à celle des saccades des quatre dernières sessions de renforcement en utilisant l'équation de Lebedev et al. (1996). La figure 35 indique que les paramètres a avant et après apprentissage sont similaires : en moyenne, il passe de 105 à 106 entre la ligne de base et les dernières sessions de renforcement (test de Wilcoxon, $p = 0.401$). Cela témoigne que notre procédure n'altère pas les propriétés dynamiques des saccades même si la variabilité de leur gain est modifiée.

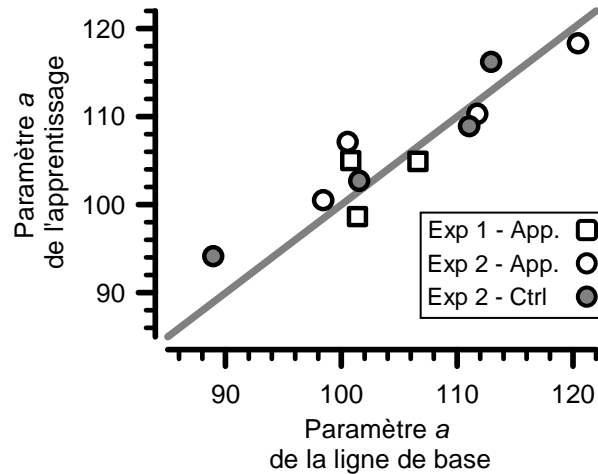


Figure 35. Paramètres a de la modélisation des *main sequences* de la ligne de base et des 800 saccades des dernières sessions d'apprentissage, obtenus via l'équation de Lebedev et al. (1996) : $\text{PicVitesse} = a\sqrt{\text{amplitude}}$. Symboles vides : participants expérimentaux, symboles gris : participants du groupe contrôle.

Par ailleurs, les théories du contrôle optimal prédisent que le caractère stéréotypé des mouvements reflèterait l'adoption par le système d'une stratégie idéale, telle qu'il minimiserait certains coûts (Körding, 2007; Wolpert & Ghahramani, 2000; Wolpert et al., 2001). En ce qui concerne les saccades oculaires, cette stratégie correspondrait à la *main sequence* (Harris, 1995; Harris & Wolpert, 1998, 2006; van Beers, 2008).

Nous avons décrit dans la première partie de la thèse que selon Harris (1995), le système oculo-moteur minimiserait les coûts liés d'une part à la durée des saccades et d'autre part à leur manque de précision. Pour cela, il réduirait le gain saccadique afin de minimiser la proportion de saccades hypermétriques (dont la durée est plus longue) et d'éviter les saccades correctives. Par conséquent, il y aurait une relation systématique entre la variabilité de la position de l'œil à la fin des saccades et le gain médian. Plus précisément, une variabilité élevée devrait s'assortir d'un gain médian plus faible que lorsque la variabilité est réduite. Qu'en est-il pour nos participants ?

Le tableau 2 résume les relations entre les changements de la variabilité et ceux du gain médian observées pour chacun d'entre eux.

Tableau 2. Changement de l'écart-type et de la médiane du gain pour chaque participant (↗ augmentation, ↗↗ augmentation supérieure à 60% pour la variabilité, et à 10% pour la médiane, ↘ diminution, ns = tendance non significative. En gris : relation prédite par le modèle de Harris, 1995).

	Participant	Ecart-type	Médiane
Exp. 1 Groupe expérimental	S1	↗↗	↘
	S2	↗↗	↘
	S3	↗↗	ns (↗)
Exp. 2 Groupe expérimental	S4	↗↗	ns (↘)
	S5	↗↗	↘
	S6	↗↗	ns (↘)
	S7	↗↗	ns (↗)
Exp. 2 Groupe contrôle	S8	↗	↘
	S9	↗	↗↗
	S10	↗	↗
	S11	↘	↗↗

Ce tableau révèle que la relation prédite par le modèle de Harris (1995) n'est observée que pour cinq participants (S1, S2, S5, S8 et S11) sur les onze. De surcroît, la relation inverse (une augmentation à la fois de la variabilité et de la médiane du gain) est constatée pour deux participants (S9 et S10). En outre, chez quatre participants expérimentaux sur sept (S3, S4, S6 et S7), les augmentations importantes de la variabilité n'ont pas été accompagnées d'un changement de la médiane du gain. Globalement, chez les participants expérimentaux, la corrélation entre les proportions de changements du gain médian et des *U Values* (calculées à partir des quatre dernières sessions d'apprentissage par rapport à la ligne de base stabilisée) n'était pas significative ($r = -0.322, p = 0.481$).

Cette absence de cohérence dans la relation entre les changements du gain et de la variabilité va à l'encontre du modèle de Harris (1995) et confirme qu'il est possible, via les contingences de renforcement où l'agent renforçateur est de nature auditive, de manipuler ces deux paramètres de manière indépendante. Autrement dit, la variabilité de l'amplitude des saccades semble relever, en partie, d'un apprentissage opérant.

IV. Méthode de l'étude 3 : renforcement de la variabilité de l'amplitude des saccades dans une tâche de recherche visuelle

Cette étude visait à tester l'hypothèse selon laquelle la variabilité de l'amplitude des saccades dépend des conséquences *visuelles* de ces saccades. En effet, si des conséquences de nature auditive permettent de renforcer différents niveaux de variabilité, comme le montre l'étude précédente, la fonction des conséquences visuelles des saccades n'a pas été directement examinée.

Les participants étaient placés dans une situation impliquant une tâche de recherche visuelle, dans laquelle ils devaient chercher une cible parmi des distracteurs.

4.1. Objectifs des deux expériences et caractéristiques des participants

La première des deux expériences de cette étude visait à vérifier que la contingence entre les mouvements oculaires et le fait de voir la cible est nécessaire pour contrôler la variabilité de l'amplitude saccadique. La cible était visible uniquement après certaines saccades, elle était affichée à la place d'un distracteur à l'issue des mouvements, en fonction des programmes de renforcement opérant. Dans la même logique que dans l'étude précédente, deux groupes de quatre personnes y ont participé. Pour un groupe expérimental (S1 à S4) le renforcement était contingent à la variabilité des amplitudes saccadiques. Pour les quatre autres participants (S4 à S8) d'un groupe contrôle *yoke*, le renforcement était non contingent : les conséquences de leurs saccades dépendaient de ce qu'avaient obtenu les participants du groupe expérimental.

Puis cinq autres personnes (S9 à S13) ont effectué une seconde expérience ayant pour objectif de tester le transfert de cet apprentissage à des situations de recherche visuelle conventionnelle – c'est-à-dire où la cible est présente parmi les distracteurs indépendamment des saccades du participant.

Ces 13 personnes, âgées de 18 à 28 ans, avaient une vision normale ou corrigée, et participaient pour la première fois à une expérience sur les mouvements oculaires. Elles étaient toutes naïves au regard du but de cette recherche. Elles ont signé un formulaire de consentement au préalable (annexe 1). Elles ont été recrutées parmi les étudiants de l'université, et

dédommagées de 50 euros pour leur participation. Cette somme était ensuite augmentée selon leurs performances à l'expérience.

Avant l'expérience, un exemple d'image comportant une cible à rechercher parmi des distracteurs leur était montré. La consigne qui leur était donnée, celle de « voir le maximum de cibles possible » (consigne qui figure en annexe 4), ne stipulait rien quant à la variabilité des comportements.

4.2. Equipement et stimulus

Les mouvements oculaires étaient mesurés à l'aide d'une caméra à infrarouges (Eyelink, 1000 Hz) commandée via l'interface du constructeur (SR Research). Ce système était connecté à un second ordinateur gérant l'enregistrement des réponses ainsi que la présentation des différents stimulus, grâce à un programme écrit sous Matlab. Les images ont été générées à l'aide de l'extension de la *Psychophysics Toolbox*. Elles étaient affichées sur un écran (Iiama HM204DT, 100 Hz) situé à 60 cm de la tête des participants stabilisée à l'aide d'une mentonnière et d'une barre frontale.

Avant chaque session, les participants devaient procéder à la calibration de l'appareil en regardant deux séries de neuf points répartis sur l'écran. En outre, entre chaque essai, soit neuf fois par session expérimentale, le programme imposait de regarder le centre de l'écran pour une recalibration en un point. Cela permettait aussi aux participants de fermer les yeux s'ils en ressentaient le besoin.

Après chaque calibration, les participants appuyaient sur une touche du clavier pour lancer l'expérience. Un point central (un carré blanc de 0.35° de côté) apparaissait sur un fond gris uniforme (luminance : 12.5 cd/m^2). A sa disparition, 750 ms plus tard, un son (de 360 Hz et d'une durée de 90 ms) était émis et l'une des images apparaissait.

Une image correspondait à une configuration de 24 items (figure 36), des cercles noirs de 0.4° de diamètre comportant en leur centre un tiret. Distracteurs et cibles se différenciaient par l'orientation de ce tiret : 45° vers la gauche pour les premiers et 45° vers la droite pour les secondes. Ils étaient suffisamment espacés (de 2.32° minimum) pour ne pas pouvoir être identifiés en vision périphérique.

De plus, les images ont été construites de manière à ce que la distribution des distances inter-items soit la plus uniforme possible. Ainsi, en ce qui concerne les distances inter-items des images d'apprentissage, la fréquence maximale était de 0.05 et la *U Value* de 0.93.

Lors de la première expérience, contrairement à la seconde, les configurations d'items de la ligne de base étaient différentes de celles des conditions de renforcement. Dans toutes les autres conditions expérimentales, huit images ont été obtenues à partir d'une seule configuration d'items ayant subi toutes les transformations de symétrie et de rotation possibles. Par ailleurs, à chaque image correspondait un masque, c'est-à-dire que les items étaient recouverts d'un carré gris qui apparaissait transitoirement, lors de chaque saccade oculaire. Ce masque avait pour but d'éviter la perception de mouvement lorsqu'une cible devait remplacer un distracteur.

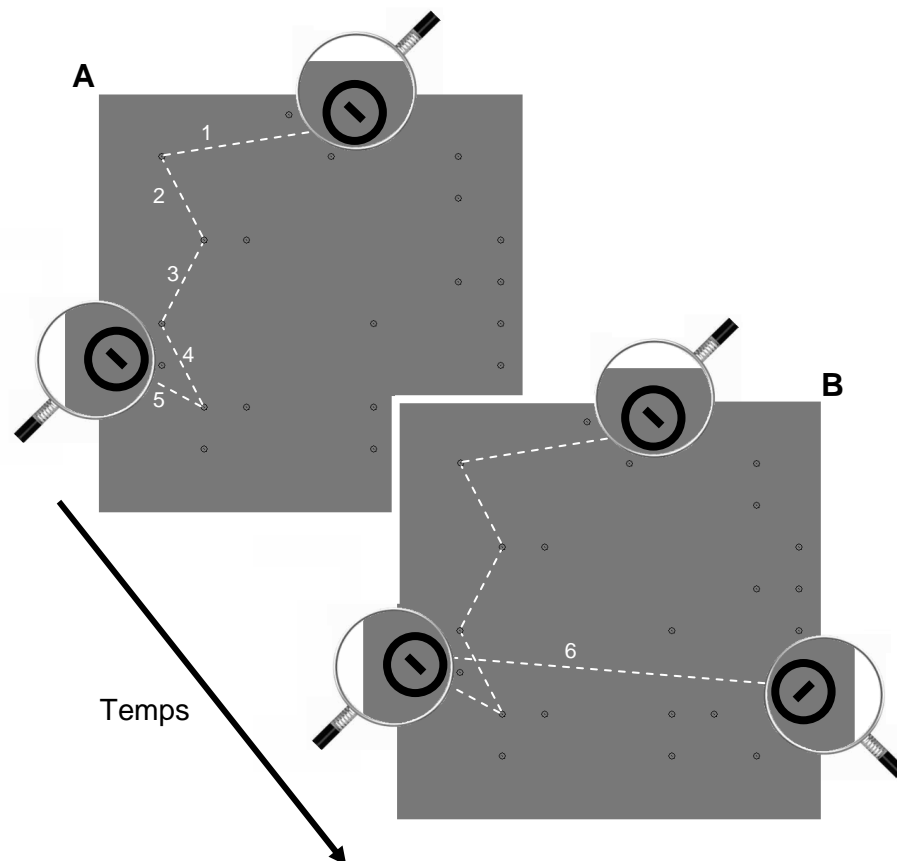


Figure 36. Représentation des stimulus utilisés dans l'étude 3. A : image avec deux distracteurs en agrandissement. B : ajout d'une cible remplaçant un distracteur après une saccade dont l'amplitude était peu fréquente (les traits en pointillés symbolisent les saccades, les chiffres indiquent le parcours des yeux).

4.3. Procédures des expériences

4.3.1. Expérience 1 : la nécessité de la contingence entre les mouvements oculaires et leurs conséquences pour contrôler la variabilité – *Yoke* control

Afin de montrer que les changements de la variabilité des amplitudes saccadiques étaient contrôlables et imputables aux contingences de renforcement, trois conditions ont été mises en place : une *ligne de base*, une condition *d'apprentissage* où un façonnage visait à augmenter progressivement la variabilité (pour le groupe expérimental), et enfin des sessions dites de *récupération* où les contingences de renforcement étaient modifiées en vue de réduire cette variabilité.

a. Ligne de base

Lors de cette première condition, les participants ont effectué 10 essais de recherche visuelle conventionnelle, consistant à trouver la cible dans chacune des images présentées – la cible occupant alors aléatoirement l'une des localisations des items, indépendamment des mouvements oculaires des observateurs. Un essai correspondait à l'affichage d'une configuration d'items particulière, distincte de celles qui étaient utilisées lors des conditions de renforcement. Dès que cette cible était trouvée, un son (de 450 Hz, d'une durée de 90 ms) était émis, puis un point de recalibration apparaissait au centre de l'écran et la configuration d'items suivante était présentée.

b. Apprentissage : façonnage d'un niveau de variabilité élevée

Dans cette condition expérimentale, les configurations d'items, parmi lesquels initialement seuls des distracteurs étaient visibles (figure 36 A), étaient présentées dans un ordre aléatoire. La cible était affichée en guise d'agent renforçateur après les saccades dont l'amplitude correspondait aux critères de renforcement (figure 36 B), accompagnée du son utilisé lors de la ligne de base.

Pour les quatre participants du groupe expérimental, la procédure de renforcement opérant utilisée était comparable à la « procédure de seuil » employée par exemple par Denney et Neuringer (1998) ainsi qu'à celle qui a été utilisée dans les sessions d'apprentissage de l'étude

précédente, si ce n'était qu'une procédure de façonnage y était couplée. Les réponses étaient renforcées si leur fréquence relative était inférieure à un seuil donné.

Les classes utilisées pour déterminer les critères de renforcement ont été obtenues à partir de la distribution des 41 distances inter-items observées – la plus petite distance étant de 2.32° et la plus grande de 22.95° . Plus précisément, ces distances ont été catégorisées en classes de 0.39° (10 pixels). Les classes ne contenant aucune distance inter-items observable étaient intégrées à la classe inférieure, si bien que les classes étaient inégales. La figure 37 illustre un exemple de calcul des critères de renforcement pour une saccade extraite de la dernière session d'apprentissage du participant S1.

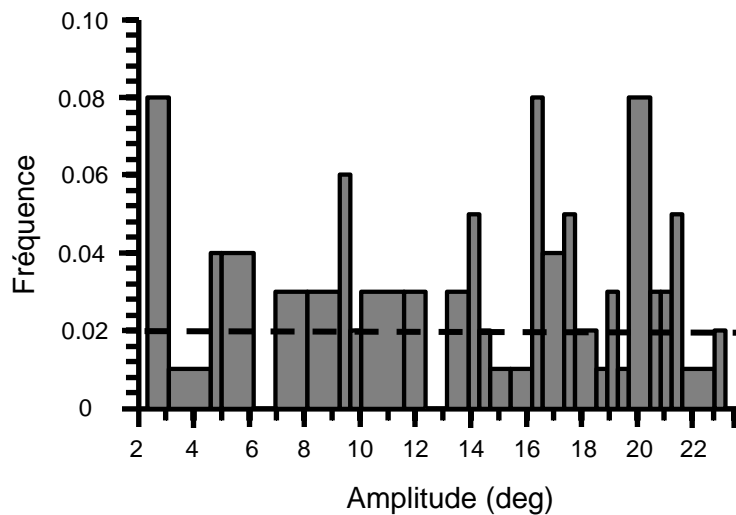


Figure 37. Critères de renforcement pour l'une des saccades de la dernière session d'apprentissage du participant S1 : fréquence relative des classes des distances inter-items observables (calculées à partir des 100 saccades précédentes) et fréquence seuil (en pointillés) en dessous de laquelle l'amplitude de la saccade allait donner lieu à l'affichage de la cible.

Dès qu'un trajet entre deux items était effectué, la distance entre ces items venait grandir la fréquence relative de la classe correspondante. Les fréquences des différentes distances inter-items parcourues étaient calculées à partir des 100 saccades les plus récentes. Pour les 100 premières saccades de chaque session, les saccades de référence correspondaient aux dernières amplitudes de la session antérieure. Pour la première saccade de la session initiale, toutes les

fréquences relatives étaient artificiellement définies comme équivalentes, la distribution évoluant ensuite peu à peu en fonction des réponses du participant.

Quant à la procédure de façonnage, elle consistait à accroître progressivement l'exigence de variabilité. Il s'agissait alors d'abaisser la fréquence seuil en dessous de laquelle l'amplitude des saccades donnait lieu à un agent renforçateur. Le seuil initial était établi à 0.05. Puis, au début de chaque session l'expérimentateur décidait de le conserver ou de le modifier en fonction de ce qu'il jugeait de la progression du participant – d'après la *U value* et le taux de renforcement obtenus lors de la session précédente.

Une session d'apprentissage était constituée de 10 essais (chacun correspondant à la présentation aléatoire de l'une des configurations d'items). Le changement d'essai avait lieu dès que 10 saccades sur les 50 précédentes avaient été renforcées. Dans ce cas, le participant obtenait 10 centimes d'euros, ce dont il était informé par l'affichage des mots « x x 0.10 € » (x correspondant au nombre de fois où cela se produisait lors de la session en cours). A l'issue de l'expérience, les participants recevaient le montant total ainsi gagné.

Pour qu'un essai soit réussi, le nombre maximum de saccades non renforcées ne devait toutefois pas dépasser 50. Dans le cas contraire, le participant devait continuer à parcourir les items jusqu'à l'émission d'une saccade correspondant aux critères de renforcement. Les mots « changement d'essai » apparaissaient alors et aucune somme d'argent n'était obtenue.

Lors d'un essai non réussi, le décompte des 10 saccades renforcées sur les 50 précédentes se reportait sur l'essai suivant. Cela pouvait entraîner un changement d'essai rapide – parfois même après l'affichage d'une ou deux cibles – accompagné du gain de 10 centimes d'euro.

Le nombre de sessions effectuées (allant de 24 à 28) dépendait du temps mis pour atteindre un niveau de performance stable et proche d'une *U-Value* de 0.80.

Quant aux quatre participants du groupe contrôle *yoke*, l'ordre d'apparition des configurations d'items, la fréquence d'affichage des cibles en vision centrale et donc la somme d'argent gagnée (soit la fréquence de renforcement non contingent) correspondaient à ce qu'avait obtenu les participants du groupe expérimental correspondant.

c. Récupération : façonnage d'une diminution de la variabilité

Par rapport à la condition précédente, le principal changement concernait le seuil servant à déterminer les critères de renforcement. Afin de réduire la variabilité, les saccades qui donnaient lieu à un agent renforçateur (pour le groupe expérimental) étaient celles dont l'amplitude était située dans les classes les plus fréquentes, c'est-à-dire les classes dont la fréquence était supérieure

à un seuil déterminé en début de session. Pour la première session de récupération, ce seuil était initialement fixé à 0.10. La procédure de façonnage consistait à élever progressivement la valeur de ce seuil.

Comme le taux de renforcement des saccades les plus fréquentes risquait d'être plus élevé que celui de la condition précédente, l'expérimentateur pouvait manipuler également la probabilité de renforcement. Cela avait pour objectif de rendre les taux de renforcement équivalents et constants tout au long de l'expérience – du moins autant que possible. Tout comme la valeur du seuil, cette probabilité était décidée au regard des données des sessions antérieures.

Le nombre de sessions effectuées (de 9 à 27) dépendait du temps requis pour obtenir une *U Value* inférieure à 0.50 sur plusieurs sessions consécutives.

De même que lors de la condition précédente, les participants du groupe contrôle ont eux aussi passé des sessions de « récupération », le renforcement étant alors encore indépendant de l'amplitude de leurs saccades. Par souci de concision, les termes « d'apprentissage » et de « récupération » seront conservés pour désigner les conditions de renforcement non contingent de ce groupe contrôle.

4.3.2. Expérience 2 : test du transfert de l'apprentissage à des situations de recherche visuelle conventionnelle

Cette expérience avait pour objectif de tester si les différents niveaux de variabilité induits par les procédures de renforcement opérant se transféraient à des situations où la cible était présente parmi les distracteurs, indépendamment des saccades des observateurs, comme dans n'importe quelle situation de recherche visuelle. Les cinq participants passaient les trois conditions expérimentales de *ligne de base*, *apprentissage* et *récupération*. Seuls les changements par rapport à l'expérience précédente sont précisés ci-dessous.

a. Ligne de base

Les 10 configurations d'items de cette ligne de base – avec une cible occupant aléatoirement l'une des localisations – étaient tirées au hasard parmi les huit configurations qui avaient été utilisées lors de l'apprentissage et de la récupération de l'expérience précédente (configurations

utilisées dans les conditions suivantes également). Aucun autre changement n'a été effectué dans celle ligne de base.

b. Apprentissage : façonnage et test d'un niveau de variabilité élevée

Les modifications de la procédure concernaient uniquement l'introduction d'essais tests à la fin de cette condition. Une fois que la performance des participants était stable, des sessions supplémentaires ont été ajoutées, pendant lesquelles des configurations d'items identiques à celles de la ligne de base (c'est-à-dire comportant une cible) étaient insérées de manière aléatoire parmi les 10 essais d'apprentissage de la session, sans que cela ne soit signalé. Dans cette condition, le nombre de sessions effectuées allait de 14 à 48.

Lors de ces essais tests, dès que les participants avaient trouvé la cible, les items disparaissaient et les participants pouvaient lire les mots « changement d'essai ». Quatre à six sessions de ce type ont été réalisées selon les participants, en fonction du nombre de saccades émises lors des essais tests avant de pouvoir trouver la cible – jusqu'à ce qu'au moins 100 saccades aient été totalisées sur l'ensemble de ces essais tests.

c. Récupération : façonnage et test d'un niveau de variabilité faible

Les essais tests de recherche visuelle conventionnelle étaient introduits de la même façon que précédemment, lors de trois à sept sessions de récupération passées une fois que la performance était stable (les participants ont effectué 8 à 51 sessions de récupération).

4.4. Acquisition et analyse des données

Tant qu'une configuration d'items était affichée, le système Eyelink mesurait *online* les mouvements oculaires. Le début et la fin des saccades étaient détectés lorsque la vitesse et l'accélération des mouvements de l'œil dépassaient les seuils respectifs de $30^\circ/\text{s}$ et $8000^\circ/\text{s}^2$. A l'issue d'une saccade dont l'amplitude correspondait aux critères de renforcement, la cible apparaissait 25 ms en moyenne après l'arrivée de l'œil sur l'item.

Pour l'examen et l'affichage *offline* des données, un programme a été élaboré afin de pouvoir passer en revue chaque saccade et exclure les éventuels clignements d'œil ou les essais bruités

malencontreusement identifiés comme des saccades. Les saccades dont la durée était supérieure à 100 ms étaient systématiquement écartées de l'analyse. De même, les éventuelles saccades intra-items (3.75% des saccades en moyenne) ainsi que les saccades amenant l'œil à plus de 2° d'un item, quel qu'il soit (3.02% des saccades en moyenne) étaient elles aussi éliminées.

La principale variable dépendante, la *U Value* (voir figure 31) a été calculée pour chaque participant et chaque condition expérimentale, par bloc de 800 saccades – ainsi que saccade par saccade, à partir des 800 saccades antérieures.

Les proportions de changements de *U Value* entre les 800 dernières saccades des différentes conditions expérimentales ont été calculées de la manière suivante :

$$\frac{U \text{ de l'apprentissage (ou récupération) - } U \text{ de ligne de base (ou apprentissage)}}{U \text{ de ligne de base (ou apprentissage)}} \times 100 \quad (7)$$

Notons que lors des essais tests de la seconde expérience, le nombre de saccades était largement inférieur à 800. Afin d'éviter que la comparaison entre les *U Values* n'en soit trop biaisée, les effectifs des différents types d'essais ont été égalisés, en ne conservant que les dernières saccades des essais dits « expérimentaux » (c'est-à-dire les essais d'apprentissage et de récupération où la cible apparaissait de manière contingente à la variabilité des amplitudes saccadiques).

Enfin, les pics de vitesse et les durées de fixation des items ont également été relevés.

De la même manière que dans l'étude précédente, les résultats des différentes conditions ont été comparés individuellement avec la méthode du Bootstrap (Efron et al., 1995), soit via la construction des intervalles de confiance (95^{ème} percentile), soit via des comparaisons par paires a priori (tests bilatéraux).

V. Résultats de l'étude 3

Chez les participants des deux groupes expérimentaux, la manipulation des contingences de renforcement a effectivement pu induire différents niveaux de variabilité dans les amplitudes saccadiques. Autrement dit, le fait de voir une cible dans une tâche de recherche visuelle a pu fonctionner comme un agent renforçateur. En effet, lorsque cette cible était visible de manière non contingente à la variabilité des saccades, les changements n'étaient pas similaires.

5.1. Expérience 1 : la nécessité de la contingence entre les mouvements oculaires et leurs conséquences pour contrôler la variabilité – *Yoke control*

La figure 38 A permet de comparer les distributions des amplitudes des 800 dernières saccades des conditions d'apprentissage et de récupération pour un participant du groupe expérimental (S1). L'abscisse de ce graphique correspond aux différentes classes de distances inter-items utilisées pour déterminer les critères de renforcement. Les deux distributions sont très différentes : tandis que celle de l'apprentissage est relativement uniforme, celle de la récupération est étalée vers la droite – les saccades les plus courtes devenaient alors les plus fréquentes. Ceci s'est reflété dans la *U Value* qui est passée de 0.57 lors de la ligne de base à 0.91 pour les 800 dernières saccades de l'apprentissage, et qui est redescendue à 0.53 à l'issue de la récupération. Dans le premier cas, cela correspondait à un changement de *U Value* de +59.22%, et dans le second, à un changement de -42.10%.

En revanche, ces changements importants n'ont pas eu lieu pour le participant S5 du groupe contrôle, dont les saccades étaient renforcées de manière non contingente à la variabilité avec les mêmes séquences de cibles que celles qui avaient été vues par le participant S1. Comme l'illustre la figure 38 B, les deux distributions des distances inter-items observées lors des 800 dernières saccades d'apprentissage et de récupération sont étalées vers la droite, le mode de celle de l'apprentissage est même supérieur à celui de la distribution de la récupération. Quant à la *U Value*, elle était de 0.57 lors de la ligne de base, de 0.55 à la fin de l'apprentissage et de 0.69 à la fin de la récupération. Les changements, respectivement de -2.85% et +23.93%, étaient non

seulement nettement inférieurs à ceux observés pour le participant S1, mais encore ils ont eu lieu dans le sens inverse.

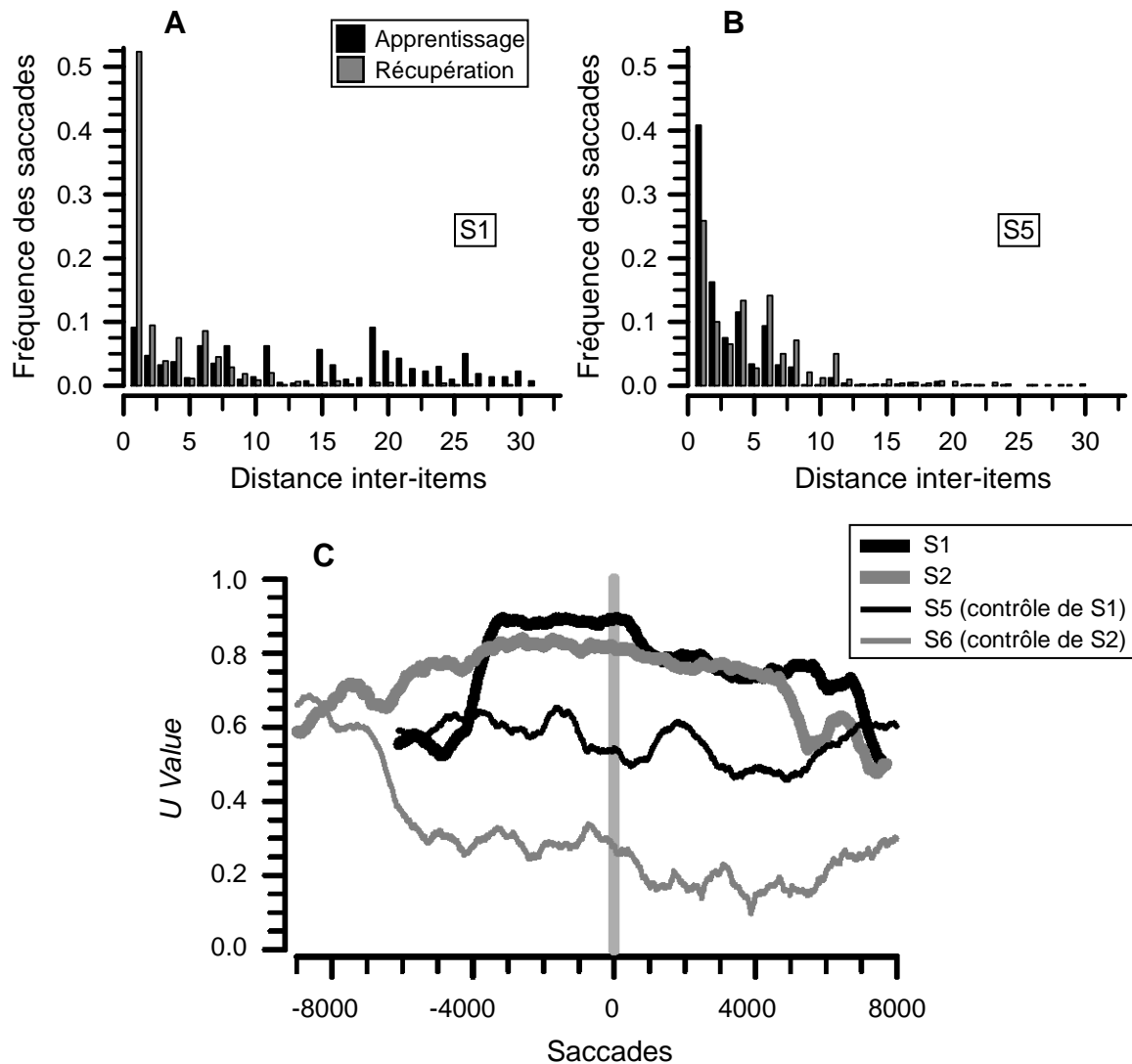


Figure 38. Résultats de l'expérience 1. A : distribution des fréquences des distances inter-items parcourues lors des 800 dernières saccades de l'apprentissage et de la récupération, pour le participant expérimental S1, et B : pour le participant du groupe contrôle correspondant, S5. C : exemples représentatifs de l'évolution de la *U Value* calculée saccade après saccade à partir des 800 saccades antérieures, pour deux participants du groupe expérimental (S1 et S2) et leur participant contrôle respectif. Les courbes sont alignées sur le changement de condition expérimentale (matérialisé par la ligne grise verticale).

La figure 38 C illustre les évolutions de la *U Value* au fil des saccades, pour deux participants du groupe expérimental (S1 et S2) et leurs participants contrôles correspondants, présentées à titre d'exemples représentatifs. En ce qui concerne les participants du premier groupe, les décours temporels des augmentations de la *U Value* étaient idiosyncrasiques, même si globalement, les niveaux de variabilité atteints à la fin des différentes conditions expérimentales étaient similaires. Quant aux participants du groupe contrôle, les changements de variabilité étaient moindres voire opposés à ceux des participants expérimentaux ; les différences inter-individuelles étaient aussi plus marquées.

La figure 39 montre que les conclusions évoquées au sujet des participants expérimentaux et contrôles précédents valaient aussi pour l'ensemble de leur groupe respectif.

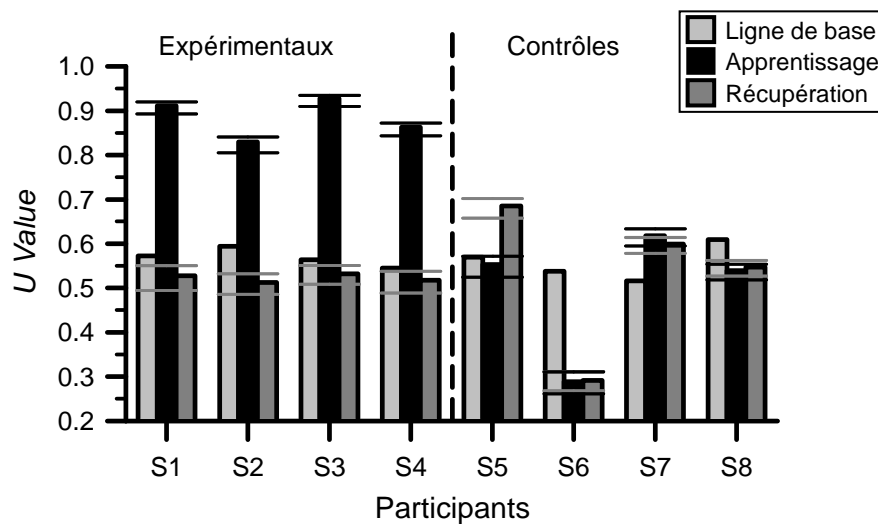


Figure 39. *U Value* pour chaque participant de l'expérience 1, calculée sur les 800 dernières saccades de chaque condition expérimentale. Les traits horizontaux correspondent aux intervalles de confiance Bootstrap (95^{ème} percentile) pour l'apprentissage (en noir) et la récupération (en gris).

Ainsi, entre la ligne de base et la fin de l'apprentissage, la proportion de changement des *U Values* pour les participants expérimentaux était de +55.48% en moyenne (écart-type = 10.83), et de -40.76% (écart-type = 2.03) entre l'apprentissage et la récupération. Si les *U Values* à la fin de l'apprentissage étaient bien sûr significativement supérieures à celles de la ligne de base, les *U Values* de la récupération étaient quant à elles significativement inférieures à ce niveau de ligne de base (la proportion de changement était de -8,05% en moyenne, écart-type = 3.96) pour les quatre participants.

En revanche, pour les participants du groupe contrôle, les proportions de changements étaient plus faibles, respectivement de -10.22% (écart-type = 30.43) et +5.77% (écart-type = 12.28) et les différences inter-individuelles plus importantes. Une augmentation de la variabilité à la fin de la première condition expérimentale n'a été observée que pour un participant sur trois (S7), et il n'y a eu aucune diminution significative lors de la seconde condition.

5.2. Expérience 2 : test du transfert de l'apprentissage à des situations de recherche visuelle conventionnelle

5.2.1. Essais expérimentaux : réplication de l'expérience précédente

La figure 40 (à gauche) indique que lors des essais expérimentaux (c'est-à-dire les essais où l'apparition de la cible était contingente à la variabilité), les changements des *U Values* étaient importants. Ils étaient comparables à ceux observés pour les participants du groupe expérimental précédent. En effet, les proportions moyennes de changement des *U Values* étaient de +54.30% (écart-type = 7.23) entre la ligne de base et l'apprentissage, et de -41.66% (écart-type = 16.62) entre la fin de l'apprentissage et la récupération.

Notons que comparativement au groupe expérimental précédent, les différences inter-individuelles observées lors de la récupération étaient plus importantes. Par rapport à la ligne de base, la *U Value* a changé de -10.12% en moyenne (écart-type = 24.86) : elle a diminué pour deux participants (S11 et S13), elle a augmenté pour S10 et est revenue au niveau de base pour S9 et S12.

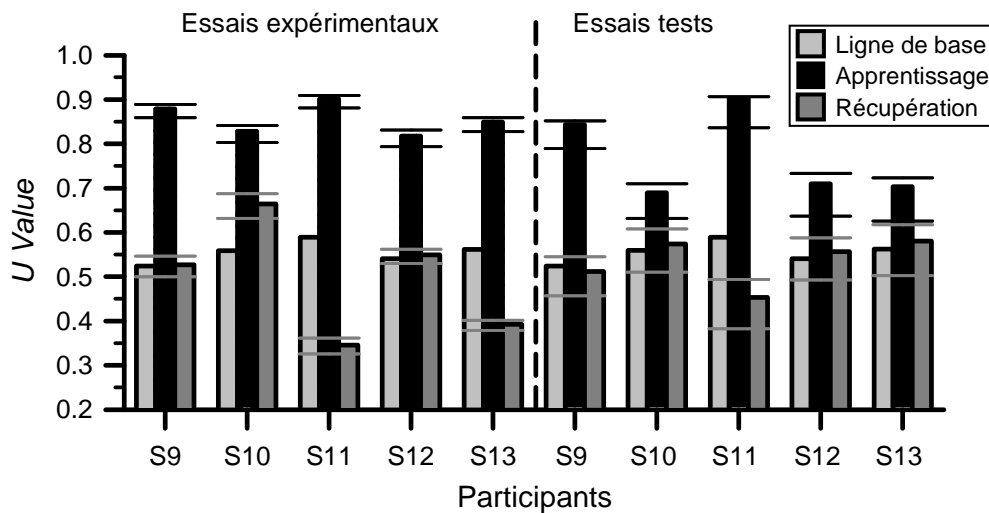


Figure 40. *U Values* pour chaque participant de l'expérience 2. A gauche, pour les essais expérimentaux (c'est-à-dire concernant les saccades émises lorsque l'apparition de la cible était contingente à la variabilité), les valeurs ont été calculées sur les 800 dernières saccades de chaque condition expérimentale. A droite, pour les essais tests, les effectifs ont été égalisés à partir de l'effectif le plus faible. Les traits horizontaux correspondent aux intervalles de confiance Bootstrap (95^{ème} percentile) pour les sessions d'apprentissage (en noir) et de récupération (en gris).

Pourrait-on y voir un effet de l'organisation spatiale des items lors de la ligne de base ? En effet, dans la première expérience, les configurations d'items présentées lors de la ligne de base étaient différentes de celles utilisées lors des conditions de renforcement, tandis que les configurations employées dans la seconde expérience étaient identiques quelle que soit la condition expérimentale. Si cela avait un effet, on pourrait s'attendre à des différences selon les expériences d'une part entre les lignes de base, et d'autre part entre les proportions de changement des *U Values* par rapport à ces lignes de base.

La figure 41 indique les résultats de cette analyse. Ni les niveaux de base, ni les proportions de changement des *U Values* des deux expériences ne différaient significativement (tests de Wilcoxon Mann-Whitney bilatéral, $p = 0.286$ et $p=0.730$, respectivement). Par ailleurs, la corrélation entre le changement des *U Values* et le niveau de variabilité de la ligne de base n'est pas significative ($r = -0.592$, $p = 0.093$, la valeur du coefficient de corrélation s'expliquant avant tout par les valeurs extrêmes). Ces éléments indiquent que la modification de la *U Value* ne

dépendait pas de la ligne de base, mais était attribuable à la manipulation des contingences de renforcement.

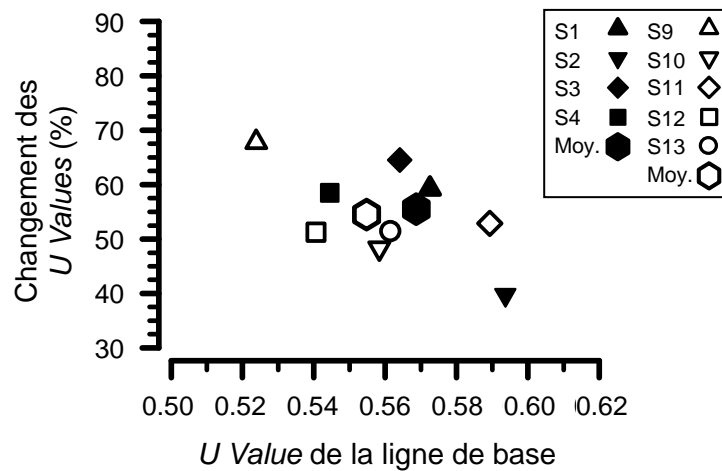


Figure 41. Proportion de changement des *U Values* (800 dernières saccades de l'apprentissage par rapport à la ligne de base) en fonction de la *U Value* de la ligne de base. Les hexagones correspondent aux moyennes de chacune des deux expériences.

5.2.2. Essais tests : transfert à des situations de recherche visuelle conventionnelle

Le principal objectif de cette seconde expérience était de tester si le niveau de variabilité induit par la manipulation des contingences de renforcement pouvait se transférer aux essais tests. Dans ces derniers, la cible était visible à une position fixe dès le début de l'essai, indépendamment des mouvements oculaires du participant.

Comme le montrent les intervalles de confiance Bootstrap représentés sur les barres à droite de la figure 40, lors des essais tests de l'apprentissage l'augmentation de la variabilité était significative pour tous les participants ; la proportion moyenne des changements était de +38.78%.

Puis, lors des essais tests de récupération, la variabilité a significativement diminué pour l'ensemble des individus ; la proportion moyenne des changements entre ces essais tests et ceux réalisés à la fin de l'apprentissage était alors de -29.02%. La *U Value* des essais tests réalisés à la fin de la récupération était équivalente à celle de la ligne de base pour tous les participants excepté S11 pour qui ce niveau de variabilité est descendu en dessous de celui de la ligne de base – un

effet cependant moins marqué que pour les essais expérimentaux. Ce sont les données de ce participant qui expliquent en grande partie la proportion de changement négative, de -3.37% en moyenne, entre la ligne de base et ces essais tests.

De manière générale, par rapport aux essais expérimentaux, les changements observés lors des essais tests étaient moins importants. En moyenne, entre les essais expérimentaux et les essais tests, les différences entre les proportions de changement étaient de 11.47% pour l'apprentissage et de 14.47% pour la récupération (différences calculées à partir d'un nombre équivalent de saccades).

C'est ce qu'illustre la figure 42. Lors de l'apprentissage (symboles pleins), la *U Value* des essais tests était soit inférieure à celle des essais expérimentaux, pour trois participants sur cinq (S10, S12 et S13, tests Bootstrap bilatéraux, $p < 0.001$), soit équivalente (pour S9 et S11, tests Bootstrap bilatéraux, $p = 0.692$ et $p = 0.293$, respectivement). De même, lors de la récupération, les changements des essais tests étaient inférieurs dans deux cas sur cinq (pour S11 et S13, tests Bootstrap bilatéraux, $p < 0.001$) ; aucun changement n'a été constaté pour S9, S10 et S12 (tests Bootstrap bilatéraux, $p = 0.913$, $p = 0.198$ et $p = 0.040$, respectivement).

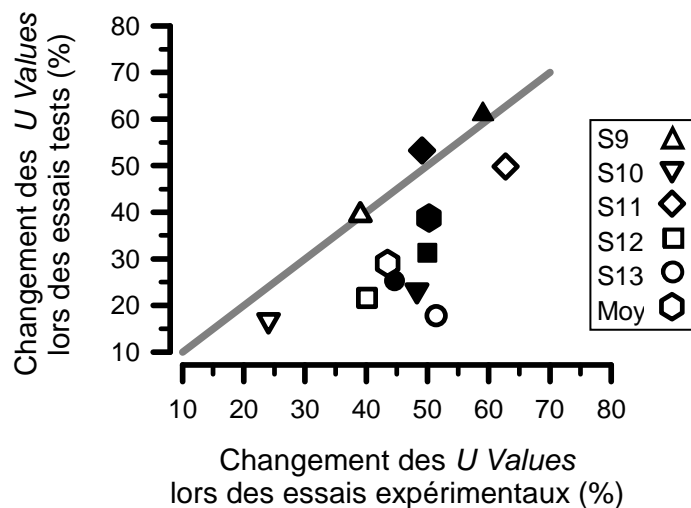


Figure 42. Proportion de changement (valeurs absolues) des *U Values* lors des essais tests par rapport à celle des essais expérimentaux, pour chaque participant de l'expérience 2. Symboles pleins : changement entre la ligne de base et l'apprentissage ; symboles vides : changement entre l'apprentissage et la récupération. Les hexagones correspondent aux moyennes de chacun des deux groupes.

Avant tout, en ce qui concerne ces essais tests, le fait que les changements entre la ligne de base et l'apprentissage et entre l'apprentissage et la récupération soient tous significatifs (figure 40 : ces changements sont au minimum de 23.22 et -16.78%, respectivement, ces deux valeurs concernant le même participant S10) indique que les modifications de la variabilité induites par les procédures de renforcement se sont transférées (au moins en partie) à des situations de recherche visuelle conventionnelle.

VI. Discussion de l'étude 3

Cette étude, qui se situe dans la continuité de la précédente, vérifie l'influence des contingences de renforcement – un renforcement de nature visuelle – sur la variabilité de l'amplitude de saccades dans une situation où les participants doivent chercher des cibles parmi des distracteurs.

La première expérience, qui permet de comparer les résultats d'un groupe expérimental pour lequel l'affichage des cibles dépend de la variabilité des saccades d'une part, et ceux d'un groupe contrôle *yoke* d'autre part, montre que la contingence entre les mouvements oculaires et leurs conséquences est nécessaire dans cet apprentissage.

Quant à la seconde expérience, elle révèle que cet apprentissage se transfère, au moins partiellement, à des situations de recherche visuelle conventionnelle (c'est-à-dire dans lesquelles une cible est présente parmi les distracteurs sans que son apparition ne dépende des mouvements de l'observateur).

En les élargissant au domaine oculo-moteur, ces résultats corroborent les conclusions qui portent sur le renforcement de la variabilité de séquences d'actions (voir Neuringer, 2002, 2009, pour des revues de littérature) telles que les appuis sur les touches d'un clavier chez l'être humain (Hopkinson & Neuringer, 2003; Neuringer, 1986; Neuringer & Voss, 1993; Souza et al., 2012).

Ils pourraient également compléter les données de Myers et Gray (2010) qui observent que lorsque les participants cherchent plusieurs fois une cible parmi des distracteurs, dans des configurations d'items sinon identiques, du moins analogues les unes aux autres, la similarité entre les trajets des saccades augmente. Selon ces auteurs, l'apprentissage (compris comme la modification des comportements liée à la simple exposition répétée à des situations identiques) réduirait la variabilité.

Nos données semblent suggérer que les stéréotypies constatées par Myers et Gray (2010) pourraient refléter le fait que dans leur expérience les contingences de renforcement ne cibleraient pas la variabilité en tant que telle. Dans cette expérience, les « contingences de renforcement » correspondraient au fait que toutes les séquences de saccades, quelles qu'elles soient, sont suivies de l'agent renforçateur : la détection de la cible.

Or le renforcement continu et non contingent à la variabilité des réponses est connu pour augmenter le caractère stéréotypé des réponses (Antonitis, 1951; Boulanger et al., 1987; Lee et al.,

2007). Dans ces conditions, répéter exactement les réponses précédentes, qui ont toutes été suivies de l'agent renforçateur, évite d'en perdre – autrement dit la variabilité a alors un coût élevé (Gharib et al., 2004). Les données de Myers et Gray (2010) pourraient refléter l'effet de ce renforcement continu.

Nos données témoignent en tout cas que la réduction de la variabilité des saccades oculaires au cours de recherches visuelles successives n'est pas systématique, mais dépend (en partie) de contingences de renforcement, de leur exigence au regard de cette dimension.

La discussion de cette étude aborde les autres effets que nos procédures de renforcement pourraient avoir sur les saccades, en dehors de ceux qui sont exclusivement liés à la contingence opérante à deux termes entre la variabilité des amplitudes saccadiques et l'agent renforçateur (l'apparition de la cible). Nous examinerons d'abord la question du contrôle discriminatif de la variabilité. Puis, nous étudierons l'influence éventuelle du taux de renforcement sur cette propriété. Enfin, nous analyserons les effets de la procédure utilisée sur les temps de fixation ainsi que sur la *main sequence*.

6.1. Mise en place de contrôles discriminatifs de la variabilité

Des expériences réalisées chez des animaux effectuant des séquences d'actions mettent en évidence le contrôle de la variabilité de ces séquences par des stimulus discriminatifs, autrement dit par des contingences à trois termes (Denney & Neuringer, 1998; Page & Neuringer, 1985). Dans notre étude, des contrôles discriminatifs donnant lieu à l'émission de différents niveaux de variabilité ont aussi pu être établis.

En premier lieu, des éléments du contexte global dans lequel se déroule l'expérience (par exemples le laboratoire, les configurations d'items, la tâche à réaliser) ont pu acquérir cette fonction de stimulus discriminatifs. Ceci, visible dans la mesure où les performances ne diminuent pas d'un jour à l'autre (comme l'indiquent notamment les évolutions des *U Values* de la figure 38), conforte la nature opérante de la variabilité des amplitudes saccadiques.

En second lieu, l'établissement d'un contrôle discriminatif pourrait être décelé à l'intérieur même des sessions tests de l'expérience 2, dans lesquelles le transfert de l'apprentissage à des situations de recherche visuelle conventionnelle était examiné.

En effet, pendant les dernières sessions (d'apprentissage et de récupération) de l'expérience 2, les participants doivent attendre l'affichage de la première cible pour pouvoir discriminer les essais tests des essais expérimentaux. Si l'essai est interrompu dès que la première cible apparaît, l'essai était un test. S'il n'y a pas de changement d'essai immédiatement après la vue de la première cible, l'essai est un essai expérimental. L'arrivée de cette première cible peut alors constituer un stimulus discriminatif annonçant le renforcement d'un niveau de variabilité très élevé (pour les sessions d'apprentissage) ou au contraire très faible (pour les sessions de récupération).

Pour voir si la procédure expérimentale installe un tel contrôle discriminatif, les saccades des essais expérimentaux ont été séparées. Il s'agit de considérer d'une part les premières saccades émises jusqu'à l'arrivée de la première cible (les participants n'ont alors pas de stimulus leur permettant de discriminer le type d'essai en cours, comme dans les essais tests), et d'autre part les saccades émises après l'arrivée de cette cible.

S'il y a un contrôle discriminatif lié à l'arrivée de la première cible, la variabilité des premières saccades ne devrait donc pas être différente de celle des essais tests. Par contre, elle devrait différer de celle des saccades émises après l'arrivée de la cible.

La figure 43 indique que les résultats des différents tests effectués ne permettent pas de conclure en faveur de cette hypothèse, les effets de ces types de saccades n'étant pas systématiques.

Les barres les plus sombres correspondent aux résultats de l'apprentissage, les plus claires, à leur droite, à ceux de la récupération. Les crochets en haut de la figure indiquent les différences significatives (comparaisons par paires a priori, tests Bootstrap bilatéraux, correction de Dunn-Sidak, $p < 0.05$) : entre les essais tests et les dernières saccades d'apprentissage d'une part (crochets larges), et entre les essais tests et les premières saccades d'apprentissage d'autre part (crochets étroits). La vérification de l'hypothèse du contrôle discriminatif est donc matérialisée par la présence d'un crochet large seul, sans crochet étroit.

En ce qui concerne les sessions d'apprentissage (crochets noirs), l'hypothèse d'un contrôle discriminatif lié à l'arrivée de la première cible n'est vérifiée que pour les participants S10 et S12. En revanche, il n'y a aucune différence significative entre les différents types de saccades pour S9 et S11. Quant à S13, les deux différences testées sont significatives.

En ce qui concerne les sessions de récupération (crochets gris), l'hypothèse du contrôle discriminatif ne s'observe que pour S11 et S13. La seule autre différence significative concerne les

essais tests et les premières saccades des sessions de récupération de S9 – une différence d'ailleurs difficilement explicable.

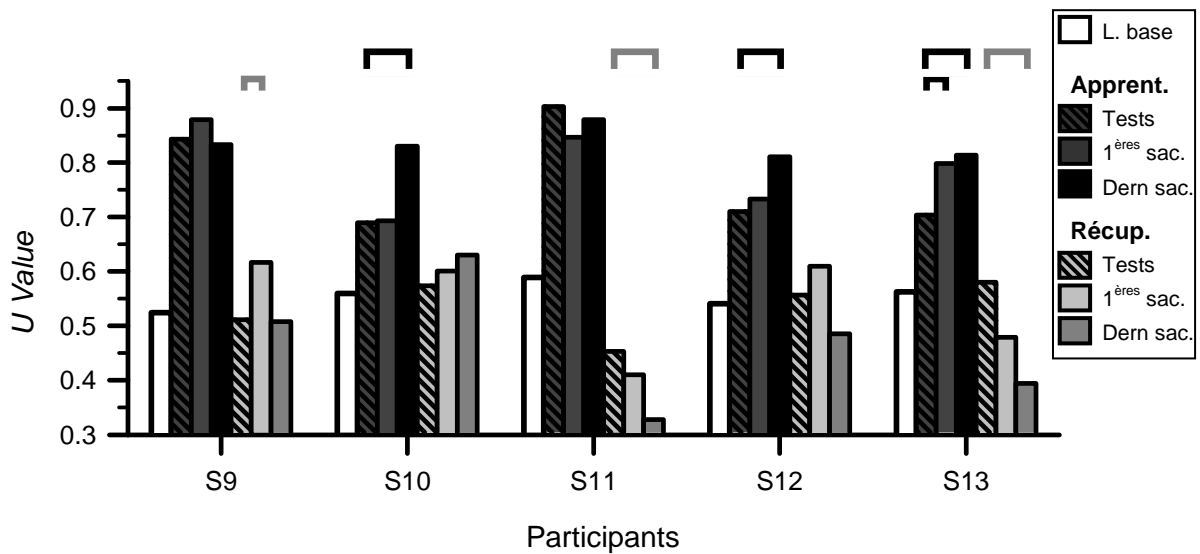


Figure 43. Analyse de l'hypothèse d'un contrôle discriminatif lié à la procédure expérimentale lors des dernières sessions tests pour chaque participant de l'expérience 2. Les trois barres foncées correspondent à l'apprentissage et les trois claires, à droite, à la récupération. Pour chacune de ces deux conditions expérimentales, les barres hachurées correspondent aux *U Values* des essais tests. Les barres centrales (gris très foncé et gris très clair) indiquent les *U Values* obtenues à partir des premières saccades (« 1^{ères} sac. ») émises jusqu'au premier affichage d'une cible lors des essais expérimentaux. Les barres plus foncées (noires et grises) à leur droite correspondent aux *U Values* calculées sur les saccades des essais expérimentaux émises après l'apparition de la première cible (« Dern sac. »). Les crochets (noirs pour les sessions d'apprentissage et gris pour la récupération) indiquent les différences significatives obtenues lors des tests Bootstrap bilatéraux (comparaisons par paires a priori, correction de Dunn-Sidak, $p < 0.05$).

En bref, le manque de cohérence des résultats ci-dessus empêche de conclure sur l'éventuelle fonction discriminative de l'affichage de la première cible. Le contexte global de l'expérience, qui permet notamment de maintenir le niveau de variabilité atteint d'une session à l'autre, semble donc avoir une plus grande influence. Il semble permettre notamment que l'apprentissage se transfère à des situations de recherche visuelle conventionnelle.

Le fait que la différence entre les *U Values* des essais tests et des essais expérimentaux ne soit visible que dans la moitié des cas (crochets larges sur la figure 43) reste toutefois inexpliquée.

6.2. Une influence indirecte des taux de renforcement ?

Dans cette troisième étude, non seulement nous n'avons pas utilisé de procédure de renforcement percentile pour contrôler précisément la probabilité de renforcement, comme Machado (1989, 1992) ou Madelain et al. (2007) l'avaient fait, mais encore la fréquence seuil déterminant les critères de renforcement était ajustée par l'expérimentateur. Par rapport à l'étude précédente, cela rend la procédure moins standardisée et risque davantage d'entraîner des fluctuations dans les taux de renforcement. Ceci est problématique dans la mesure où, selon plusieurs études portant sur des séquences d'actions effectuées par des animaux, la fréquence de renforcement peut avoir un effet sur la variabilité comportementale (Lee et al., 2007; Neuringer, 2002, 2009) et entrer en interaction avec les programmes de renforcement direct de la variabilité (Grunow & Neuringer, 2002, Machado, 1989, 1992).

Certes les résultats finaux du groupe contrôle *joke* permettent de conclure que la contingence est nécessaire pour modifier la variabilité. Néanmoins, les différences inter-individuelles relevées à la fin des sessions expérimentales, que ce soit dans les groupes expérimentaux dont les participants sont tous soumis aux mêmes contingences de renforcement ou bien dans le groupe contrôle, incitent à analyser l'influence du taux de renforcement plus en détails.

La figure 44 présente à titre d'exemples les évolutions de la *U Value* et du taux de renforcement au fil des sessions pour le participant S1 et le participant contrôle correspondant, S5.

L'un des éléments marquants concerne l'augmentation brusque de la *U Value* à la huitième session, qui s'accompagne d'une forte montée du taux de renforcement, tandis que lors de la récupération, c'est la diminution rapide de la *U Value* lors de la session 48 qui coïncide avec un taux de renforcement élevé. Cela est lié au fait que le renforcement dépend de ces niveaux de variabilité respectifs, autrement dit, à la contingence de renforcement.

En fait, cette figure souligne avant tout l'impossibilité d'établir un lien univoque entre taux de renforcement et *U Value* : le taux de renforcement, qui – de manière regrettable – n'est pas stable, dépend à la fois du comportement des participants expérimentaux et des manipulations de l'expérimentateur. En effet, les taux de renforcement plus faibles entre les sessions 8 et 48

peuvent s'expliquer par les ajustements de la fréquence seuil déterminant les critères de renforcement.

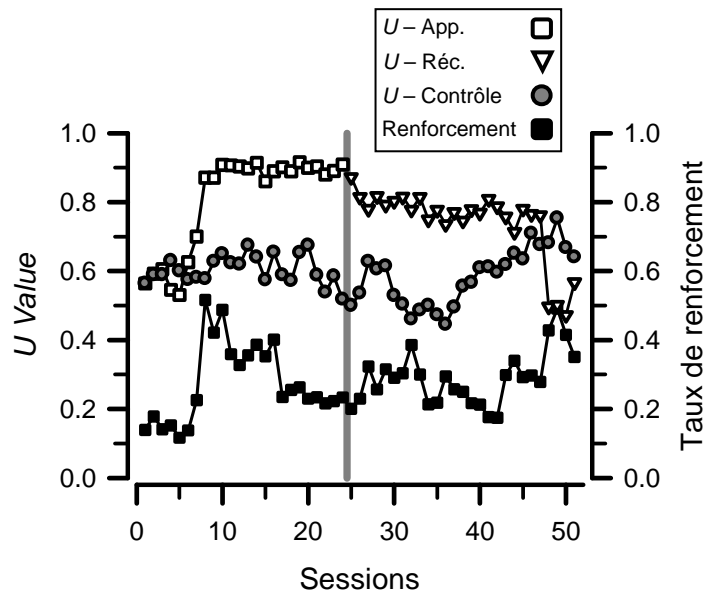


Figure 44. Exemples d'évolutions de la *U Value* et du taux de renforcement au fil des sessions pour le participants S1 (expérience 1) et le participant contrôle S5 correspondant. La ligne grise correspond au changement de condition expérimentale. (App. : apprentissage, Réc.= récupération).

Par ailleurs, les données du participant contrôle attestent que les modifications du taux de renforcement, alors non contingent à la variabilité, ne s'accompagnent pas des changements de variabilité prédits par la littérature, selon laquelle plus la probabilité de renforcement est faible, plus la variabilité augmente (Antonitis, 1951; Cherot et al., 1996; Gharib et al., 2004; Neuringer et al., 2001; Stahlman et al., 2010) : les coefficients de corrélation entre les deux variables ne sont pas significatives, comme l'indique le tableau 3.

Tableau 3. Coefficients de corrélation entre la *U value* et le taux de renforcement des différentes sessions, pour l'apprentissage et la récupération. En gris : corrélation correspondant à l'hypothèse selon laquelle la raréfaction du renforcement accroît la variabilité – et/ou que la diminution de la *U Value* entraîne une augmentation du taux de renforcement lorsque le renforcement vise à diminuer la variabilité.

	Participant	<i>r</i> apprentissage	<i>p</i>	<i>r</i> récupération	<i>p</i>
Exp. 1 Groupe Expérimental	S1	0.677	< 0.001	-0.763	< 0.001
	S2	0.794	< 0.001	-0.571	0.003
	S3	0.686	< 0.001	-0.658	0.054
	S4	0.723	< 0.001	-0.713	< 0.001
Exp. 1 Groupe Contrôle	S5	0.418	0.042	0.281	0.156
	S6	-0.379	0.047	0.457	0.027
	S7	0.606	0.001	0.477	0.195
	S8	0.072	0.715	-0.6727	< 0.001
Exp. 2 Groupe Expérimental	S9	0.284	0.325	-0.318	0.087
	S10	0.808	< 0.001	0.729	< 0.001
	S11	0.807	< 0.001	0.020	0.963
	S12	0.213	0.159	-0.488	0.002
	S13	0.009	0.966	-0.590	< 0.001

Ce tableau montre qu'en ce qui concerne les participants des groupes expérimentaux, lors de l'apprentissage, aucune corrélation n'est négative : l'hypothèse selon laquelle la raréfaction du renforcement accroît la variabilité n'est donc pas vérifiée. Pour deux tiers des participants, les coefficients sont positifs : comme pour le participant S1, cette corrélation reflète la *contingence* de renforcement.

Quant à la récupération, dans les groupes expérimentaux, le coefficient de corrélation entre la *U value* et le taux de renforcement est négatif dans les deux tiers des cas, mais ceci est toujours conforme à la contingence – il est impossible de conclure précisément en terme d'intermittence du renforcement puisque les deux effets vont ici dans le même sens.

Enfin, les données des participants contrôles, pour lesquels seulement un quart des corrélations sont négatives, suggèrent que dans cette étude l'intermittence du renforcement n'exerce pas une influence prépondérante sur la variabilité comportementale, celle-ci dépendant bien avant tout des contingences de renforcement.

6.3. Les effets de la procédure de renforcement opérant sur les durées de fixation

De même que dans l'étude précédente, les latences sont susceptibles d'être affectées par la procédure de renforcement de la variabilité. L'examen de cette propriété pourrait éventuellement donner des indications sur l'utilisation de stratégies cognitives liées à la sélection des cibles. Ainsi, Stritzke et al. (2009) et Schütz et Gegenfurtner (2010), dans leurs études sur les saccades émises en direction de certaines régions d'une cible associée à des gains monétaires, observent que les performances sont corrélées positivement avec les latences. Selon eux, cela reflèterait l'intervention de processus cognitifs – de stratégies – liés au traitement des informations relatives aux récompenses.

Notons que dans cette étude 3, il est impossible de déterminer la latence exacte d'une saccade par rapport à l'apparition d'une cible en vision périphérique puisque les items sont continuellement présents. C'est donc la durée de fixation des items qui est étudié ici.

La figure 45 trace la proportion de changement de la *U Value* des sessions expérimentales (c'est-à-dire des sessions d'apprentissage et de récupération) en fonction du changement de ces durées de fixations. A la recherche d'indices relatifs à l'utilisation de stratégies, nous avons aussi analysé ces durées selon que les saccades ont été suivies de l'affichage d'une cible ou non.

Pour tous les participants, la durée de fixation lors des sessions expérimentales a augmenté par rapport à celle de la ligne de base (elle a augmenté de 54.48% en moyenne, test de Wilcoxon bilatéral, $p < 0.001$) : comme le montre la figure 45 : la majorité des points, que ce soit ceux des groupes expérimentaux ou ceux du groupe contrôle, se situent à droite de la ligne verticale matérialisant une absence de changement des durées de fixation.

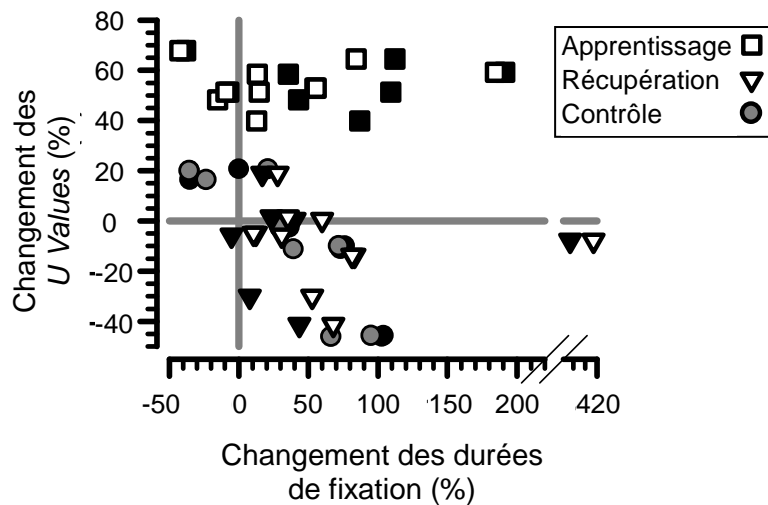


Figure 45. Proportion de changement des U Values en fonction de la proportion de changement des durées de fixation des items pour chaque participant de l'étude 3. Ces proportions ont été obtenues à partir des 800 dernières saccades de l'apprentissage (carrés) et de la récupération (triangles), et des sessions correspondantes pour les participants du groupe contrôle (cercles). Les symboles noirs correspondent aux données des saccades renforcées par une cible, les symboles gris et blancs à celles des saccades non suivies de l'affichage de cette cible.

On ne peut écarter l'effet de la fatigue puisque la ligne de base est mesurée au tout début de l'expérience, alors que les saccades des sessions expérimentales sont obtenues à la fin d'une séance quotidienne de 45 à 60 minutes.

Cette augmentation des latences pourrait refléter également l'utilisation de stratégies. Par exemple, Souza et al. (2012), dans leur recherche relative à l'influence des instructions et des règles que se forment les individus sur la variabilité de séquences d'appuis sur des touches d'un clavier, observent des latences différentes selon que les individus doivent se comporter de manière systématique ou aléatoire. Un autre élément arguant en faveur de ces stratégies concerne l'existence de certains changements brusques de la variabilité, observés par exemple chez le participant S1 (figure 38 C).

Par ailleurs, pour les participants des groupes expérimentaux, lors de l'apprentissage, la durée de fixation précédant les saccades renforcées par l'apparition de la cible est plus élevée que celle

précédant les saccades non renforcées par la cible (572 versus 461 ms, respectivement, test de Wilcoxon bilatéral, $p = 0.008$). Par contre, pour la récupération, l'inverse est constaté : les saccades renforcées par la cible sont précédées des durées de fixation les plus courtes (604 versus 684 ms, test de Wilcoxon bilatéral, $p = 0.020$).

Cela pourrait s'expliquer à la lumière de l'expérience de Sohn et Lee (2006) dans laquelle la dernière saccade émise à la fin de séquences répétitives est renforcée. Ils observent que les latences de ces derniers mouvements sont plus courtes que celle des autres saccades. Cet effet est probablement dû au stimulus discriminatif que constitue la position de la dernière cible de la série.

Dans notre expérience, lors des sessions d'apprentissage, les participants pouvaient difficilement associer une localisation particulière à l'apparition d'une cible. En revanche, lors de la récupération, le renforcement de la répétition des saccades de mêmes amplitudes pouvait se traduire par l'association – accidentelle – d'une localisation précise et d'une cible. La situation est alors plus similaire à celle étudiée par Sohn et Lee (2006) – elle peut refléter l'existence d'un contrôle discriminatif.

Il semble difficile de tirer d'autres inférences au sujet de l'existence éventuelles de règles, de stratégies.

Neuringer et al. (2001), qui étudient la variabilité de séquences d'appuis sur des leviers chez des rats, notent que les intervalles inter-réponses sont corrélés positivement avec le niveau de variabilité. Mais ceci ne se retrouve pas dans nos données : les coefficients de corrélation entre les changements des durées de fixation et des *U Values* ne sont pas significatifs (pour l'apprentissage, $r = 0.044$, $p = 0.911$; pour la récupération, $r = -0.045$, $p = 0.911$). Quant au groupe contrôle, c'est la relation inverse qui est observée : plus la durée de fixation augmente, plus la variabilité diminue ($r = -0.903$, $p = 0.002$).

Si l'on ne peut écarter l'implication de stratégies, ces dernières ne peuvent toutefois pas rendre compte à elles seules des résultats. A l'instar d'autres recherches réalisées sur la variabilité opérante chez l'être humain (Hopkinson & Neuringer, 2003; Neuringer, 1986; Neuringer & Voss, 1993; Souza et al., 2012), il est peu probable que sans les contingences de renforcement, les changements des *U Values* des participants expérimentaux soient si importants – du moins cela resterait à tester.

De plus, l'utilisation de stratégies n'explique pas pourquoi l'ampleur des changements de variabilité est similaire quel que soit le participant des groupes expérimentaux, et ce malgré des décours temporels idiosyncrasiques (comme le suggèrent les exemples donnés par la figure 38 C).

Même si l'existence de stratégies propres à chaque participant ne peut être exclue, cela ne remet nullement en cause la nature opérante de la variabilité comportementale.

6.4. Analyse de la *main sequence*

La procédure de renforcement de l'étude 2 n'a pas altéré les propriétés dynamiques des saccades, ce qui contredit les données de notre première étude ainsi que celles d'autres auteurs qui ont renforcé différenciellement certaines saccades et ont observé une augmentation du pic de vitesse – que le renforcement soit contingent à cette propriété (Madelain & Paeye, 2011) ou non (Takikawa et al., 2002; Xu-Wilson et al., 2009). De même que précédemment, les *main sequences* ont été analysées par le biais de l'équation de Lebedev et al. (1996). La figure 46 permet de comparer les paramètres obtenus à partir des 800 dernières saccades des sessions expérimentales par rapport à celles de la ligne de base pour chacun des participants.

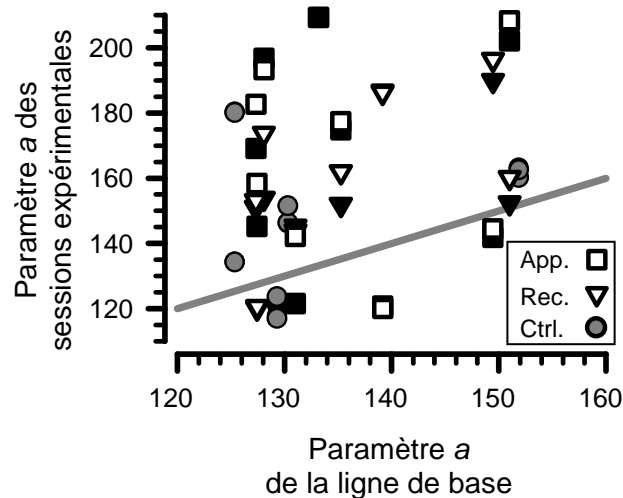


Figure 46. Paramètres a de la modélisation des *main sequences* (via l'équation de Lebedev et al. (1996) : $\text{PicVitesse} = a\sqrt{\text{amplitude}}$) de l'apprentissage (carrés) et de la récupération (triangles) en fonction du paramètre a de la ligne de base. Les cercles indiquent les données du groupe contrôle. Les symboles noirs correspondent aux données des saccades renforcées, les symboles gris et blancs à celles des saccades non renforcées.

Cette figure 46 révèle que de manière générale, les paramètres a de la ligne de base, compris entre 125 et 152 (écart-type = 9.60) varient moins que ceux des sessions expérimentales (qui sont compris entre 117 et 221, écart-type = 29.06). Les valeurs de ce paramètre diffèrent de celles des études précédentes parce qu'elles dépendent du mode de calcul du pic de vitesse. Dans cette troisième étude, les mesures des pics de vitesse utilisées sont celles qui sont fournies directement par l'algorithme élaboré par le constructeur de l'Eyelink (SR Research). Une vérification a montré que ces mesures sont très fortement corrélées avec celles que l'on obtient *offline* en recherchant le maximum de la vitesse enregistrée tout au long de la saccade – l'examen d'une session (la ligne de base du participant S10) indique une corrélation de 0.979 ($p < 0.001$) entre les deux mesures de pics de vitesse, avec une moyenne de 262°/s pour les données *online* versus 207°/s pour les données *offline*. Dans ce dernier cas, le paramètre a de l'équation de la *main sequence* est réduit de 24%, et il est cohérent avec les valeurs obtenues dans les études précédentes.

Pour l'ensemble des participants, les valeurs de a sont globalement plus élevées lors des sessions expérimentales que lors de la ligne de base (161 versus 135 en moyenne, test de Wilcoxon bilatéral, $p = 0.002$), ce qui traduit une augmentation des pics de vitesse lors des sessions expérimentales.

De plus, pour les participants expérimentaux, les paramètres a sont également plus élevés pour les saccades non suivies de la cible que pour les saccades qui en sont suivies, mais cette différence est très faible (172 versus 165 en moyenne, test de Wilcoxon bilatéral, $p = 0.039$). Cela ne s'observe pas pour les participants du groupe contrôle (pour lequel $a = 147$ dans les deux cas, tous les cercles noirs correspondant aux saccades renforcées étant dissimulés par les cercles des saccades non renforcées). Enfin, aucune différence n'est constatée entre les paramètres a de la *main sequence* des sessions d'apprentissage et de récupération (168 versus 166 en moyenne, test de Wilcoxon bilatéral, $p = 0.820$) pour les participants des groupes expérimentaux.

L'effet le plus notable est donc l'importante augmentation des pics de vitesse par rapport à la ligne de base. Cela conteste une éventuelle influence de la fatigue, et confirme qu'il est possible, via le renforcement différentiel des saccades, d'altérer cette relation soi-disant stéréotypée qu'est la *main sequence*.

VII. Conclusion des études 2 & 3 : le contrôle de la variabilité de l'amplitude saccadique par les contingences de renforcement

Ces deux études confirment qu'en manipulant les contingences de renforcement, il est possible d'induire expérimentalement différents niveaux de variabilité oculo-motrice, que ce soit la variabilité du gain de saccades isolées ou celle des amplitudes de saccades intégrées dans des séquences. Cette conception de la variabilité des saccades oculaires comme relevant (en partie) d'un apprentissage opérant se démarque d'un certain nombre de théories courantes.

Ainsi nos données montrent que la variabilité ne correspond pas uniquement aux conséquences délétères d'un bruit interne incontrôlable. Nos travaux se démarquent de certaines approches actuelles du contrôle moteur optimal qui s'appuient sur les régularités des paramètres des mouvements (comme la *main sequence*) pour inférer l'existence de mécanismes visant à générer les mouvements les plus efficaces possibles, notamment en minimisant la variabilité (Harris, 1995; Harris & Wolpert, 1998, 2006; van Beers, 2008)

Par ailleurs, notre approche se distingue de celle de modèles computationnels qui ont récemment été proposés pour rendre compte des séquences de saccades émises dans des situations de recherche visuelle. L'un des plus pertinents à ce jour est celui élaboré par Geisler et ses collègues (Geisler, Perry, & Najemnik, 2006; Najemnik & Geisler, 2005, 2008), qui prennent en compte la sensibilité inégale de la rétine dans leurs calculs bayésiens. Selon leur modèle, les séquences de saccades optimales minimisent le nombre de mouvements nécessaires pour localiser la cible et maximisent l'information obtenue à l'issue de chaque saccade. Ce type de théories se focalise donc sur les régularités (comme le nombre moyen de saccades par essai ou le fait que les distributions des amplitudes saccadiques émises lors des essais de recherche visuelle sont étalées vers la droite avec un mode compris entre 1 et 5° – ce qui correspond d'ailleurs aux distributions de la ligne de base de notre étude). La variabilité inter et intra-individuelle, pourtant importante, n'est pas expliquée, mais elle est masquée par le moyennage d'un nombre important de données.

Loin de concevoir la variabilité uniquement comme le reflet d'écarts par rapport à une moyenne ou un comportement optimal, nos résultats mettent l'accent sur le contrôle que peuvent exercer les conséquences fonctionnelles des saccades sur cette propriété. Ainsi, des conséquences

arbitraires, comme un son, ou plus naturelles, comme le fait d'apercevoir une cible parmi des distracteurs, peuvent amener non pas à réduire, mais à augmenter la variabilité à l'issue d'un apprentissage – à condition que ces conséquences soient contingentes à cette propriété.

Ces études s'inscrivent dans la continuité de celle de Madelain et al. (2007) qui montrent que la dispersion et la médiane des latences saccadiques peuvent être manipulées indépendamment l'une de l'autre via l'application de programmes de renforcement opérant. Elles apportent également des arguments supplémentaires en faveur de la nature opérante des mouvements oculaires (Madelain, Paeye, & Darcheville, 2011).

DISCUSSION & PERSPECTIVES

Les travaux présentés dans cette thèse montrent qu'il est possible de contrôler expérimentalement les distributions des amplitudes des saccades oculaires en manipulant les contingences de renforcement. Les conséquences visuelles immédiates des saccades ont été modifiées afin de pouvoir faire suivre d'un agent renforçateur uniquement les mouvements qui répondent aux critères de renforcement en vigueur.

La première étude montre que le gain de saccades horizontales peut être ainsi progressivement diminué ou augmenté, en renforçant les saccades de plus en plus courtes ou de plus en plus longues.

La deuxième indique que la variabilité de l'amplitude de ces saccades est elle aussi une dimension opérante, c'est-à-dire une propriété contrôlée par ses propres conséquences. En renforçant d'abord les saccades dont les amplitudes sont les moins fréquentes, la dispersion des gains saccadiques peut être augmentée ; puis, en renforçant les saccades dont les amplitudes sont les plus fréquentes, la variabilité revient à son niveau de base. Dans cette étude, les agents renforçateurs sont de nature auditive.

La troisième recherche atteste que des conséquences visuelles peuvent aussi fonctionner comme agent renforçateur dans des procédures opérantes visant à manipuler la variabilité de l'amplitude des saccades. Les participants ont pour tâche de trouver une cible parmi des distracteurs ; la cible est affichée à la place d'un distracteur en guise d'agent renforçateur, c'est-à-dire de manière contingente à la variabilité croissante ou décroissante des saccades. Ces procédures entraînent respectivement soit une augmentation, soit une diminution de cette propriété.

Ces recherches mettent en exergue le contrôle de l'amplitude saccadique, y compris de sa variabilité, par des conséquences *fonctionnelles*, quelle que soit leur nature (auditive et/ou visuelle). Ces conséquences sont définies par le fait qu'elles déterminent l'apparition et le maintien des comportements (Skinner, 1953). Autrement dit, le contrôle et la morphogenèse au sein du système oculo-moteur semblent donc relever (en partie) d'un mécanisme général d'apprentissage – un apprentissage opérant (Madelain, Paeye, & Darcheville, 2011).

En ce sens, cette approche se distingue de la conception computationnelle du contrôle et de l'apprentissage oculo-moteurs. En effet, les théories classiques postulent l'implication de mécanismes spécialisés dans des processus de prédiction et de comparaison de signaux permettant la recalibration de gains internes (Bahcall & Kowler, 2000; Chen-Harris et al., 2008; Collins & Wallman, 2011; Ethier et al., 2008; Wong & Shelhamer, 2011). Elles avancent également que le système oculo-moteur apprendrait à optimiser des fonctions de coûts,

notamment ceux liés à la durée et à la variabilité des saccades (Harris & Wolpert, 1998, 2006; Karniel, 2011; van Beers, 2008; Wolpert & Ghahramani, 2000; Wolpert et al., 2001).

En ce qui concerne la variabilité, les résultats des recherches présentées dans cette thèse, cohérents avec ceux de l'analyse expérimentale du comportement (Lee et al., 2007; Neuringer, 2002, 2009), montrent qu'elle ne correspond pas uniquement aux conséquences délétères d'un bruit interne, conséquences qu'il s'agirait nécessairement de minimiser.

I. La morphogenèse au sein du système oculo-moteur, un processus de sélection comportementale

Concevoir le contrôle du gain des saccades comme relevant d'un apprentissage opérant sous-entend l'existence d'un processus de sélection comportementale. En d'autres termes, les saccades qui sont suivies d'une conséquence bénéfique pour l'organisme – en l'occurrence la vision nette d'une cible – seraient réitérées, au détriment des saccades qui ne sont pas suivies de cette conséquence.

1.1. De l'origine des variations comportementales

Tout comme la sélection naturelle, la sélection comportementale repose sur l'existence de variations (Donahoe, 2003; Donahoe et al., 1993; Skinner, 1981). La procédure opérante de façonnage est un bon exemple de l'interaction entre ces variations – nécessaires pour que des approximations du comportement cible apparaissent – et le renforcement différentiel de ces approximations, celles-ci étant alors sélectionnées au détriment des formes comportementales plus lointaines du comportement cible.

Par conséquent, on peut faire l'hypothèse que dans notre première étude, la capacité des procédures opérantes – ainsi que des procédures de double saut – d'augmenter ou de diminuer le gain des saccades repose aussi sur l'existence de variabilité dans le gain saccadique.

Skinner (1953) ou Galbicka (1994) mentionnent deux principaux mécanismes qui garantiraient ces variations comportementales dans le façonnage. Le premier est l'extinction (voir Antonitis, 1951; Neuringer et al., 2001), et le second la « généralisation de la réponse ». Cette dernière est souvent définie comme le fait que le renforcement d'un comportement possédant une forme particulière augmente la probabilité d'apparition d'autres réponses ayant la même fonction, mais des topographies différentes. Ce phénomène de « généralisation de la réponse » pourrait être lié à l'effet de la fréquence du renforcement sur la variabilité (Cherot et al., 1996; Gharib et al., 2004; Lee et al., 2007; Stahlman et al., 2010) et/ou à la présence de contingences de renforcement opérant ciblant cette dimension. C'est ce que suggèrent les études sur la variabilité

de séquences d'actions (ex : Machado, 1989, 1992; Neuringer, 2002, 2009) ou celle de différentes propriétés de réponses (Blough, 1966; Ross & Neuringer, 2002) – y compris, comme cette thèse vient de le présenter, la variabilité des saccades oculaires (voir aussi Madelain et al., 2007).

Cela signifie qu'antérieurement, des associations ont été apprises. Les analystes du comportement mentionnent en effet souvent le rôle que joue l'histoire des contingences de renforcement sur la variabilité, et ce même si ces contingences ne portent pas sur cette dimension (Lattal & Neef, 1996). Par exemple, des rats qui apprennent à appuyer sur un levier selon des exigences précises (comme n'appuyer qu'avec la patte droite) qui se relâchent progressivement (si bien que finalement n'importe quelle manière d'appuyer sur le levier donne lieu à un agent renforçateur) émettent en fin d'apprentissage des réponses dont la forme reste plus variable que ce qui est observé chez des rats pour qui les contraintes de topographie deviennent au contraire de plus en plus sévères (Stokes, 1995).

En outre, lorsque des êtres humains apprennent, via des programmes de renforcement opérant contingent à la variabilité, à diversifier leurs séquences d'appuis sur les touches d'un clavier, ils maintiennent souvent un niveau de variabilité élevé même lors de tests où celle-ci n'est plus requise par le programme de renforcement (Maes, 2003; Miller & Neuringer, 2000; Stokes & Balsam, 2001). Joyce et Chase (1990), Maes (2003), ou encore Souza et al. (2012) attribuent en partie ce maintien à l'influence d'antécédents verbaux (instructions ou règles auto-imposées), jouant le rôle de stimulus discriminatifs pour l'émission d'un certain niveau de variabilité.

Mais comment expliquer l'origine de la variabilité comportementale au niveau ontogénétique ?

Neuringer suppose l'existence d'un « générateur de variabilité interne » (un « *random generator* »), qui serait inné et s'ajusterait, se réglerait au simple contact des contingences de renforcement en vigueur, expliquant la capacité des individus à produire différents niveaux de variabilité (Neuringer, 2002; Neuringer & Jensen, 2010; Page & Neuringer, 1985). L'existence de ce mécanisme interne est toutefois difficilement vérifiable.

A l'instar de Donahoe et Palmer (1994) on peut invoquer, en dehors de facteurs tels que la variabilité liée au bruit neuronal ou à la présence de troubles physiologiques, l'influence de variables environnementales qui modifient les réflexes des nouveaux-nés. Les auteurs mentionnent notamment l'implication de l'arousal : la simple apparition de stimulations intenses ou répétées, même en dehors de contingences de renforcement, peut entraîner l'augmentation de la vigueur des réponses (Killeen et al., 1978). Selon Donahoe et Palmer (1994), l'état de « déprivation » (le fait que l'organisme n'ait pas été en contact avec certains stimulus autant que

s'il y avait eu un accès illimité) serait un autre facteur augmentant la force des réponses. Cela concernerait non seulement les réactions à ce stimulus, mais aussi celles qui sont provoquées par d'autres sources de stimulations. Donahoe et Palmer (1994) expliquent cela en terme de divergence comportementale : le stimulus dont aurait été privé l'organisme, quand il arrive enfin, entraînerait de plus grands changements que lorsqu'il est continuellement présent. Les auteurs font l'hypothèse que, suite à ces différentes variations de comportements, des processus de renforcement viendraient sélectionner des niveaux précis de variabilité, qui seraient placés également sous contrôle discriminatif (Donahoe, 2003; Donahoe & Palmer, 1994).

1.2. Renforcement direct de la variabilité opérante & façonnage : prolongement des recherches actuelles

1.2.1. Un effet facilitateur du renforcement direct de la variabilité opérante ?

Accroître la variabilité comportementale signifie modifier et éventuellement augmenter la probabilité d'émettre des réponses proches du comportement cible. Puisque le renforcement opérant peut contrôler cette variabilité tout en évitant la diminution du taux de réponses, on peut faire l'hypothèse que l'introduction de ce renforcement direct de la variabilité avant ou pendant la mise en place de façonnages pourrait faciliter les apprentissages (Neuringer, 2002, 2009; Stokes, 1995). C'est-à-dire que cela pourrait rendre possible ou accélérer l'apparition de certaines formes comportementales, tandis que sans un tel procédé, cette apparition pourrait s'avérer plus lente voire impossible.

Plusieurs études qui portent sur l'apprentissage de séquences d'actions testent directement cette hypothèse, chez les animaux et chez les êtres humains. Différents groupes d'individus sont comparés : l'un reçoit un renforcement contingent à la variabilité des séquences, un deuxième groupe obtient des agents renforçateurs de manière non contingente à cette dimension, et un groupe contrôle n'obtient aucune conséquence relativement à cette propriété. Pour chaque groupe, l'émission d'une séquence cible est simultanément renforcée.

Les études de Neuringer (1993) et Neuringer, Deiss et Olson, (2000) menées chez des pigeons et des rats mettent en évidence le fait que les groupes dans lesquels le renforcement est contingent à la variabilité sont les seuls qui parviennent à émettre toutes les séquences cibles, y

compris celles qui ont initialement une faible probabilité d'occurrence. Dans leur recherche de 2002, Grunow et Neuringer renforcent quatre niveaux de variabilité différents (dans des groupes distincts) avant de mettre en place le renforcement d'une séquence cible ; ils constatent que la capacité et la vitesse d'apprentissage d'une nouvelle séquence sont corrélées positivement avec le niveau de variabilité préalablement renforcé.

Le principal problème de ces études est que lors de la ligne de base, la fréquence d'émission des séquences cibles n'est jamais nulle ; dans les expériences de Grunow et Neuringer (2002) et Neuringer et al. (2000), elle est même corrélée avec le niveau de variabilité initial des différents groupes. Par conséquent, il est difficile de dissocier l'éventuel rôle facilitateur de la variabilité et l'effet du renforcement lui-même – qui est d'accroître la probabilité d'émission des comportements auxquels il est contingent.

Maes et Van der Goot (2006) testent l'hypothèse de la facilitation de l'apprentissage par le renforcement opérant de la variabilité chez des humains. Contrairement aux recherches chez les animaux, leurs résultats indiquent que seuls les participants du groupe contrôle émettent les séquences cibles, même celles qui auparavant n'apparaissaient dans aucun des groupes (voir aussi Bizo et Doolan, 2008, pour une étude préliminaire avec des résultats similaires). Maes et Van der Goot (2006) font l'hypothèse que les individus recevant des agents renforçateurs à la suite de séquences différentes de la séquence cible ne détecteraient pas la contingence relative à cette dernière – éventuellement à cause des règles que les participants se donnent.

Le risque d'observer cette interférence est d'autant plus fort que le taux de renforcement des séquences variables est élevé : en effet, les études d'Abreu-Rodrigues et al. (2005) et Neuringer (1992) montrent dans ce cas une préférence pour l'émission de séquences variables plutôt que pour la répétition.

L'un des prolongements de cette thèse pourrait donc consister à étudier l'effet du renforcement opérant de la variabilité du gain saccadique (comme dans notre deuxième étude) sur l'apprentissage de saccades oculaires. Cet apprentissage pourrait consister à émettre des saccades dont le gain correspondrait à une valeur cible définie par l'expérimentateur (comme dans notre première étude). De même que dans la recherche de Grunow et Neuringer (2002), plusieurs groupes pourraient être constitués, pour lesquels différents niveaux de variabilité seraient renforcés.

L'utilisation d'une procédure de renforcement percentile serait alors primordiale : elle permettrait de réduire progressivement la fréquence du renforcement contingent à la variabilité tout en maintenant intact le niveau de cette dernière. En effet, les critères d'exigence des

programmes pourraient alors être ajustés de manière à compenser les effets éventuels de cette intermittence. De cette façon, la comparaison entre les groupes ne souffrirait pas d'un biais lié à la différence des taux de renforcement puisque ceux-ci pourraient être égalisés.

Il est d'ailleurs regrettable que cette procédure de renforcement percentile n'ait pas été appliquée dans nos études sur le renforcement de la variabilité. En effet, la maîtrise des effets de la fréquence de renforcement aurait peut-être permis d'atténuer certaines différences inter-individuelles. En outre, il est possible que les taux de renforcement, parfois fluctuants, y aient interféré avec les effets du renforcement opérant, direct, de la variabilité.

Premièrement, prolonger ainsi les recherches présentées dans cette thèse étendrait les recherches sur l'apprentissage de séquences d'actions au domaine sensori-moteur. Deuxièmement, cela devrait permettre de tester l'hypothèse de l'effet facilitateur de la variabilité sur l'apprentissage chez l'être humain, en évitant l'interférence entre les programmes de renforcement de la variabilité et le renforcement du nouveau comportement (Bizo & Doolan 2008; Maes & Van der Goot, 2006). Troisièmement, cela offrirait un moyen de tester la relation entre différents niveaux de variabilité *opérante* et l'apprentissage tout en évitant les problèmes de la fréquence initiale d'émission du comportement cible (rencontrés dans les études de Grunow et Neuringer, 2002; Neuringer et al., 2000), puisque les saccades à apprendre ont très peu de chances d'être émises en dehors de cette situation expérimentale.

1.3. Du développement du contrôle des mouvements oculaires

Par ailleurs, cette approche opérante du contrôle des saccades peut nous amener à formuler une hypothèse à propos du développement de ces mouvements, une hypothèse qui se distingue de celles avancées par les théories classiques du contrôle moteur.

Par exemple, Harris (1995) suggère qu'au début de la vie, les structures corticales impliquées dans la correspondance entre l'excentricité de la cible et les commandes motrices permettant à l'œil d'atteindre cette cible ne seraient pas fonctionnelles. Il se base sur les études rapportant que le gain saccadique des enfants de quelques mois est très faible et très variable, au point que les enfants réalisent souvent d'une à cinq saccades lorsqu'ils dirigent leur regard vers une cible, que celle-ci soit située à 10 ou 40° du point de fixation initial (Aslin & Salapatek, 1975). Selon Harris (1995), le coût lié à la durée du mouvement pèserait alors davantage que le coût lié à la précision dans le contrôle saccadique et le système minimiserait en priorité le temps passé dans les saccades. Ce ne serait qu'à l'issue de la maturation de ces structures que le système pourrait établir

cette correspondance entre les signaux visuels et les commandes motrices. Alors seulement, le coût lié à la nécessité de la précision des mouvements serait pris en compte.

Plutôt que d'affirmer que le contrôle du gain des saccades relèverait d'un processus de maturation corticale, nos études nous amènent à supposer que l'augmentation du gain et la réduction de la variabilité des saccades qui ont lieu lors des premiers mois de la vie pourraient résulter d'un processus d'apprentissage opérant. Cet apprentissage se produirait parce que les conséquences visuelles des saccades seraient modifiées. En effet, à la naissance, la rétine des jeunes enfants n'est pas mature (Yuodelis & Hendrickson, 1986). Les interactions visuelles limitées. Dès lors, si le gain saccadique est faible et la variabilité élevée, cela n'entraînerait pas de conséquences plus délétères que si le gain était proche de un et la variabilité faible – ce qui ne serait plus valable plus tard, par exemple lors de la lecture, pendant laquelle les mouvements saccadiques sont beaucoup plus contraints. Initialement, aucune sélection comportementale n'aurait donc lieu.

Puis, progressivement, la vision des cibles, de plus en plus claire, renforcerait différenciellement certaines formes de mouvements oculaires tandis que les mouvements amenant la fovéa trop loin de l'image des cibles seraient éteints. Autrement dit, se produirait alors une divergence comportementale, ces changements dans les actions en cours subséquents à l'apparition de stimulus susceptibles de fonctionner comme des agents renforçateurs (Donahoe et al., 1993; Donahoe & Palmer, 1994).

Cette hypothèse semblerait s'appliquer également aux mouvements de poursuite oculaire, émis lorsqu'une cible se déplace lentement. Lorsque la trajectoire des yeux suit celle de la cible, l'image de cette dernière peut être continuellement projetée sur la fovéa, et la vision de cette cible reste nette pendant le déplacement.

Le contrôle de la poursuite oculaire par les contingences de renforcement est mis en évidence notamment dans des recherches où ces mouvements oculaires sont renforcés par l'émission d'un son, diffusé de manière continue en fonction de la vitesse du déplacement des yeux. Ainsi, dans leur étude de 2003, Madelain et Krauzlis demandent à des participants de suivre une cible, alors qu'elle disparaît transitoirement (donc en l'absence de signaux visuels). Le renforcement positif permet d'augmenter progressivement la vitesse des mouvements et de réduire le nombre de saccades. Ces modifications ne s'observent pas dans un groupe *yoke*. Dans le même ordre d'idée, dans une étude préliminaire réalisée dans notre laboratoire, des participants parviennent à initier des mouvements oculaires lisses à une vitesse déterminée par l'expérimentateur, mais alors que

l'image de la cible est stabilisée sur la rétine (Paeye, Darcheville, & Madelain, 2007). Il n'y a donc pas de signal visuel provenant du déplacement de la cible.

Enfin, dans leur étude chez des nouveaux-nés, Darcheville, Madelain, Buquet, Charlier et Miossec (1999) diffusent de la musique de manière contingente aux mouvements de poursuite oculaire. En modifiant les conséquences de ces derniers, les auteurs créaient donc un phénomène de divergence comportementale : c'est ainsi qu'ils expliquent l'émission de plus en plus fréquente de la poursuite lisse par des enfants âgés de quelques jours.

En bref, nous pouvons formuler l'hypothèse que le développement ontogénétique de certaines formes de mouvements oculaires – ceux qui sont suivis de la vision nette des cibles et qui favorisent ainsi les interactions avec l'environnement, l'accomplissement de comportements visuellement guidés – résulterait d'un changement dans les contingences de renforcement.

II. Intégration de la « valeur » des conséquences des saccades dans les modèles computationnels du contrôle moteur

2.1. L'effet de la manipulation des conséquences des saccades sur les différentes propriétés de ces mouvements : un consensus ?

Suite aux nombreuses recherches utilisant le paradigme de double saut ainsi qu'à nos propres expériences, l'influence de la modification des conséquences *visuelles* post-saccadiques (le fait que la saccade soit suivie ou non de la projection de l'image de la cible sur la fovéa) sur le gain des saccades n'est plus à démontrer.

Cette assertion est assez générale pour être acceptée à la fois par l'analyse du comportement, qui envisage cette stimulation comme un agent renforçateur, et par les théories classiques du contrôle moteur, pour lesquelles le signal guidant l'adaptation saccadique ne peut être élaboré sans cette conséquence.

Cependant, malgré le nombre important d'études consacrées à ce phénomène, il persiste aujourd'hui encore quelques désaccords quant à l'influence des conséquences visuelles sur les propriétés des saccades. L'un d'eux porte sur l'effet de l'adaptation saccadique sur la *main sequence*. Les revues de littérature de Hopp et Fuchs (2004) et celle de Pélisson et al. (2010) insistent sur les résultats contradictoires obtenus lors des différentes répliques de ce paradigme.

A titre d'exemples, dans les études menées chez l'être humain, Frens et van Opstal (1994) n'ont observé aucune différence dans les *main sequences* des saccades avant et après l'adaptation induite par le paradigme de double saut ; il en va de même pour les adaptations conventionnelles de notre première étude. Collins, Semroud, Orriols et Doré-Mazars (2008) observent une altération du rapport entre l'accélération et la décélération, mais sans modification de la *main sequence* proprement dite. Alahyane et Pélisson (2005) constatent quant à eux une augmentation de la durée des saccades après l'adaptation (pour des amplitudes constantes), néanmoins ils retrouvent le même effet après l'émission de saccades contrôles sans saut intra-saccadique : la fatigue semble donc jouer un rôle dans la modification de cette *main sequence*. En revanche selon

Ethier et al. (2008), cet effet de la fatigue sur la *main sequence* serait moins important que celui du saut intra-saccadique lui-même, mais uniquement si ce dernier est négatif.

Il est donc difficile de conclure. A fortiori, répliquer simplement la mesure de saccades met en exergue une importante variabilité intra-individuelle. En effet, Bollen et al. (1993) déterminent les *main sequences* de saccades émises par 58 participants vers des cibles situées à des excentricités comprises entre 3 et 20°, à deux semaines d'intervalles. Si la corrélation entre l'amplitude et le pic de vitesse est forte quel que soit le moment où elle est mesurée, le coefficient de corrélation entre les pentes des régressions des deux *main sequences* n'est que de 0.23. Dès lors, il peut paraître surprenant que des auteurs, tels que Ethier et al. (2008), basent leur démonstration sur le changement de la *main sequence* lors des expériences de double saut.

Un second désaccord concerne l'effet de l'adaptation saccadique sur les latences saccadiques. Ainsi, aucun changement de latences n'a été observé après l'adaptation saccadique via la procédure de double saut chez les participants des expériences d'Albano et King (1989) ou Frens et van Opstal (1994), de même que chez ceux de notre première étude. Ethier et al. (2008) observent par contre une augmentation.

Ces incohérences dans les résultats sont à souligner dans la mesure où elles peuvent rendre parfois difficile l'interprétation des données que nous avons obtenues suite à l'application des procédures de renforcement opérant. Des expériences complémentaires semblent nécessaires pour pouvoir trancher ces questions, où seraient contrôlés les facteurs tels que la fatigue, les protocoles expérimentaux ou les méthodes d'analyse employées, voire l'âge des participants, celui-ci pouvant avoir aussi une influence sur les différentes propriétés des saccades (Irving et al., 2006).

Pourrait-on parler davantage d'un consensus autour de la proposition selon laquelle la « valeur » des conséquences des saccades en affecterait les propriétés ?

De prime abord, c'est la conclusion qui semble pouvoir être tirée des recherches mentionnées dans cette thèse. Pour résumer, l'ampleur et/ou la probabilité de la récompense – du jus de fruits chez le singe (Platt & Glimcher, 1999; Schroeder & Holland, 1969; Sugrue et al. 2004) ou des gains monétaires chez les êtres humains (Schütz & Gegenfurtner, 2010; Stritzke et al., 2009) – influencent la direction des saccades. Les latences sont affectées par la présentation de liquide après certaines saccades chez le singe (ex : Hikosaka et al., 2006; Lauwereyns et al., 2002; Nakamura & Hikosaka, 2006; Takikawa et al., 2002; Watanabe et al., 2003). Chez l'humain, elles sont réduites par la présence d'items à discriminer (Montagnini & Chelazzi, 2005) et les pics de vitesse diffèrent selon le type d'images affichées après les mouvements (Xu-Wilson et al., 2009).

Les expériences réalisées dans notre laboratoire ont elles aussi manipulé les conséquences des saccades : des conséquences auditives et/ou visuelles (versus aucune conséquence) apparaissent en fonction des latences (Madelain et al., 2007), des pics de vitesse (Madelain & Paeye, 2011) ou, dans les études présentées précédemment, en fonction des amplitudes et de leur variabilité, et elles permettent de modifier ces propriétés.

Néanmoins, selon la perspective théorique adoptée, la notion même de « valeur » des conséquences, elle, fait débat.

2.2. La « valeur » des conséquences saccadiques : un rôle central, des définitions différentes selon les approches théoriques

2.2.1. Une « valeur » déterminée a posteriori selon l'analyse expérimentale du comportement, une « valeur » subjective selon les théories de la prise de décision

Rappelons que la théorie de l'apprentissage opérant définit la valeur de la conséquence d'un comportement selon les effets de ce stimulus sur le comportement, sa capacité à le modifier. Cette valeur (renforçante) ne peut donc être déterminée qu'a posteriori (Francisco et al., 2008; Hodos & Kalman, 1963; Skinner, 1953).

Ainsi, Berlyne (1972) conclut que des patterns visuels complexes ont une valeur plus élevée que des patterns simples seulement après avoir testé l'effet de leur présentation contingente à l'émission d'un comportement simple (l'appui sur un bouton) sur la fréquence des réponses. L'utilisation d'un programme de renforcement concurrent permet de contrôler l'effet éventuel de l'arousal. Quant aux études réalisées dans le cadre de cette thèse, elles indiquent que les conséquences visuelles (versus auditives) ont assez de valeur pour modifier les propriétés saccadiques telles que le gain ou la variabilité des amplitudes. Cette définition de la valeur (renforçante) d'une conséquence n'implique aucun calcul interne.

Les théories de la prise de décision considèrent quant à elles la valeur des conséquences d'un comportement comme un concept d'ordre subjectif. Cette valeur est opérationnalisée par les récompenses, qui peuvent varier par exemple en termes de quantité et/ou de probabilité, ou encore de délai d'obtention. L'attribution d'une valeur à différentes options serait une étape importante dans la prise de décision, dans la mesure où elle déterminerait les objectifs des

organismes, les choix effectués parmi les différentes actions possibles, sachant que ces organismes auraient pour but général de maximiser les récompenses (ex : Glimcher, 2003; Körding, 2007).

Selon cette approche, la valeur (subjective) des conséquences des actions serait estimée suite à des calculs de probabilités effectués par le système. Ces calculs consisteraient notamment à comparer les coûts et les bénéfices associés à chaque option. Le système intégrerait également l'incertitude de ses propres estimations relatives aux conséquences (qui peuvent être très variables par exemple) de ses choix, de ses actions.

Quant à l'apprentissage, il pourrait donc être défini comme les changements comportementaux observés suite à la réévaluation de ces conséquences (via des calculs internes) – quand par exemple le système est face à une différence entre la conséquence actuelle et la conséquence attendue (Körding et al., 2007).

2.2.2. Une « valeur » subjective qui expliquerait le choix des cibles regardées ?

Ce concept de valeur (subjective) des conséquences des mouvements a été intégré dans les modèles de certains théoriciens de la prise de décision pour rendre compte du contrôle et de l'apprentissage des saccades oculaires (Glimcher, 2003; Trommershäuser et al., 2009). Selon eux, les caractéristiques visuelles des stimulus, telle la saillance, ne suffisent pas pour expliquer l'orientation des saccades (Schütz, Braun, & Gegenfurtner, 2011; Tatler, Hayhoe, Land, & Ballard, 2011).

Au niveau comportemental, l'attribution de la valeur des cibles se traduirait par exemple par le fait que la fréquence relative des saccades émises en direction de certaines cibles dépend de la probabilité ou de l'importance des récompenses qui y sont associées – en suivant la loi du *matching* (Platt & Glimcher, 1999; Sugrue et al., 2004). Cette évaluation expliquerait aussi que, lorsque les participants doivent faire des saccades en direction d'une cible associée à des gains monétaires et jouxtant un distracteur associé à des pertes (Schütz et al., 2010 ; Stritzke et al. 2009) le point d'arrivée de l'œil à l'intérieur de la cible s'éloigne de ce distracteur.

Après avoir manipulé à la fois le montant des gains et des pertes ainsi que la saillance de leurs items, Stritzke et al. (2009) et Schütz et Gegenfurtner, (2010) proposent un modèle optimal qui rend compte des données en pondérant ces deux types d'information avec comme contrainte de maximiser le gain (voir aussi Navalpakkam, Koch, Rangel, & Perona, 2010, pour des conclusions similaires). Selon ces auteurs, le système oculo-moteur, à partir de tels calculs, construirait une

représentation topographique de la valeur des items, représentation sur laquelle il se baserait pour guider ses mouvements.

En ce qui concerne la recherche visuelle, le modèle proposé par Geisler et son équipe (Geisler et al., 2006; Najemnik & Geisler, 2005, 2008) ne semble pas incompatible avec cette idée. Dans leurs études, la tâche consiste à trouver une cible présentée dans du bruit visuel. Les auteurs déterminent d'abord une « carte de visibilité » (c'est-à-dire la représentation de la probabilité que la cible soit présente à chaque localisation rétinienne) à partir de fonctions psychométriques relevées au préalable pour différentes cibles, présentées à plusieurs endroits et en fonction de plusieurs contrastes. Cette carte détermine ensuite les prédictions du modèle. Lors de la simulation de la recherche visuelle, elle est réactualisée avec l'information obtenue après chaque saccade ; les déplacements ultérieurs sont alors guidés par ce nouveau calcul.

Le modèle de Najemnik et Geisler a été élaboré de telle sorte qu'il minimise le nombre de saccades émises pour trouver la cible : les auteurs montrent qu'alors la stratégie optimale – et qui correspond le mieux aux données – consiste à diriger les saccades vers les endroits qui permettent au système d'obtenir le maximum d'informations à l'issue de chaque déplacement.

Même si les auteurs, ne manipulant pas les conséquences des mouvements, ne concluent pas explicitement en termes de récompenses, ces informations pourraient être envisagées comme telles – le système assignerait alors une valeur aux différentes localisations en fonction de la probabilité d'y collecter ces informations (Schütz et al., 2011; Trommershäuser et al., 2009).

Mais le modèle de Najemnik et Geisler se heurte à quelques limites. Ainsi, il sous-estime les amplitudes des saccades ainsi que l'augmentation du nombre de fixations lorsque les cibles sont éloignées du centre de l'écran. Il ne rend pas bien compte non plus de l'existence d'un biais en faveur des saccades horizontales (Geisler et al., 2006; Najemnik & Geisler, 2008).

De surcroît, Morvan et Maloney (2012) ont récemment testé l'hypothèse selon laquelle le système oculo-moteur se baserait effectivement sur l'optimisation des récompenses pour orienter les saccades. Dans une tâche de discrimination, leurs participants doivent effectuer une saccade unique vers une localisation parmi trois. Ces localisations, équidistantes, forment une ligne horizontale et sont situées à une distance variant aléatoirement entre 8 et 24° à chaque essai. L'item à discriminer apparaît aléatoirement soit à droite, soit à gauche, de manière équiprobable. Etant donné la baisse de la sensibilité de la rétine avec l'excentricité des cibles, pour augmenter la probabilité de répondre correctement (c'est-à-dire pour maximiser le gain), la stratégie optimale consiste à regarder au centre lorsque les éléments sont rapprochés (la vision périphérique permettant alors de discriminer les items de chaque côté). Au contraire, lorsque les éléments sont

éloignés, regarder l'une des deux extrémités de l'alignement devient plus rentable. Or pour tous les participants, la probabilité de regarder au centre et aux extrémités reste constante quel que soit l'espacement des localisations.

Morvan et Maloney (2012) en concluent que, dans cette tâche de recherche visuelle simplifiée, la sélection de la cible ne relèverait pas d'une évaluation de la probabilité de distinguer l'item, de telle sorte que le système maximiserait les informations recueillies (et donc les chances de réussite) à chaque saccade. Mais Morvan et Maloney (2012) n'avancent aucune explication de leurs résultats.

Du point de vue de l'analyse du comportement, les données de Morvan et Maloney (2012) pourraient s'expliquer par le fait que l'espacement des localisations ne constituerait pas un stimulus discriminatif contrôlant le choix des réponses (c'est-à-dire incitant à regarder au centre versus à la périphérie).

Cette interprétation s'appuie sur les données mêmes de Morvan et Maloney (2012). Dans une dernière condition expérimentale, un signal indique aux participants quelle localisation regarder pour maximiser la probabilité de répondre correctement. Le gain des participants – équivalent au taux de renforcement – n'est alors augmenté que de 9% en moyenne par rapport à la condition précédente. Il se peut que cette différence ne soit pas assez importante pour avoir été détectée et avoir entraîné un changement dans le choix opéré par les participants. Autrement dit, ce choix pourrait dépendre avant tout de contingences qui auraient été apprises rapidement lors des premiers essais, et qui pourraient favoriser accidentellement l'une des options.

Laissant de côté les concepts de « cartes de visibilité », de « représentations topographiques de la valeur des cibles », et de calculs menant à l'élaboration de stratégies optimales, nous proposons donc une interprétation en termes de stimulus discriminatifs qui signalent quelles sont les contingences de renforcement en vigueur et donc augmentent la probabilité d'apparition de certaines saccades. Cette interprétation peut, comme cela a été présenté dans la première partie de cette thèse, rendre compte de l'orientation des saccades considérées individuellement – et donc expliquer les résultats de Platt et Glimcher (1999); Sugrue et al. (2004) ou de Navalpakkam et al. (2010), Schütz et Gegenfurtner, (2010), Stritzke et al. (2009) par exemple. Il semble qu'elle puisse aussi rendre compte de l'orientation de saccades à un niveau plus global. Dans le même ordre d'idée, ces contingences à trois termes expliqueraient aussi que, dans notre troisième étude, les participants apprennent progressivement à modifier leurs séquences de saccades sans diminution de leur performance d'un jour à l'autre.

Comment falsifier l'hypothèse selon laquelle le système oculo-moteur orienterait les saccades en fonction de la valeur attribuée à différentes localisations ? Nos résultats semblent indiquer qu'il se baserait sur les propriétés des *mouvements* antérieurs – ceux qui ont été suivis de l'apparition de la cible. Néanmoins, dans nos études, les mouvements oculaires restent contraints par la localisation des différents items (c'est le cas aussi dans l'expérience de Morvan et Maloney, 2012). Par exemple, dans notre troisième étude, émettre des saccades de taille différente ou identique pourrait s'appuyer sur des stratégies consistant à associer certains items à l'apparition de la cible. C'est ce que pourrait suggérer d'ailleurs l'analyse des latences, plus courtes lors des sessions de récupération que lors de l'apprentissage et de la ligne de base.

Il pourrait être intéressant d'exploiter le paradigme de recherche visuelle dans du bruit de Najemnik et Geisler (2005), cette fois en manipulant les conséquences des saccades. La cible pourrait être présentée en guise d'agent renforçateur de manière contingente à certaines propriétés des mouvements (par exemple les saccades d'une amplitude donnée). Cela devrait avoir des répercussions sur les séquences de saccades tout en limitant la possibilité d'établir des associations entre des localisations spécifiques et l'apparition de la cible (d'autant plus si les critères de renforcement le prenaient en considération). Si cela n'élimine pas la possibilité de l'attribution de valeurs à des mouvements, cela pourrait du moins mettre au défi les explications en terme de « représentation topographique » de valeurs.

2.2.3. Une « valeur » subjective qui expliquerait les propriétés des saccades ?

Par ailleurs, le concept de valeur (subjective) des conséquences des mouvements est également intégré dans certains modèles du contrôle optimal pour rendre compte des modifications des propriétés saccadiques.

Dans une étude récente, Shadmehr et al. (2010) réexaminent le modèle proposé par Harris et Wolpert (1998, 2006). Shadmehr et al. (2010) avancent que le système oculo-moteur optimiserait des fonctions de coûts tout en maximisant les récompenses obtenues à l'issue des saccades. Dans leur modèle, l'un de ces coûts est lié à la nécessité de la précision des mouvements. Shadmehr et al. (2010) introduisent également le coût de l'effort, dont la fonction, comme la précédente, diminuerait avec la durée des saccades.

Enfin, là où Harris et Wolpert (2006) postulent un coût lié à la durée du mouvement (et donc selon eux d'une vision altérée) qui augmenterait de manière linéaire, pour Shadmehr et al. (2010) cette fonction serait hyperbolique (figure 47) car elle proviendrait de la dévaluation des

récompenses au fil du temps. Ce postulat se base sur des études comportementales qui montrent que les choix dirigés vers une récompense sont inversement corrélés avec le délai d'arrivée de ce stimulus, un phénomène bien décrit par une fonction hyperbolique (Myerson & Green, 1995). La récompense consiste alors selon Shadmehr et al. (2010) à voir la cible en vision centrale.

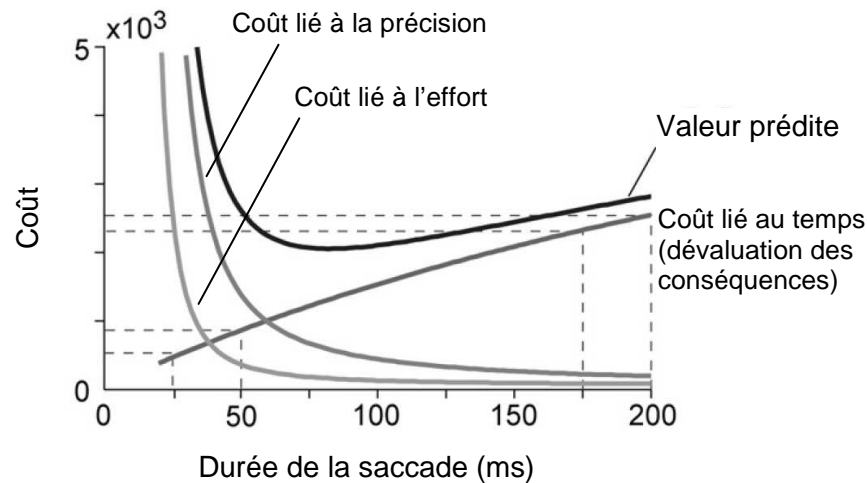


Figure 47. Simulation du coût d'une saccade de 20° en fonction de sa durée. La durée optimale de cette saccade, 85 ms, correspond au minimum de la fonction de la « valeur prédite » – qui combine les trois autres coûts liés à la nécessité de la précision, à l'effort et à la dévaluation des conséquences (source : Shadmehr et al., 2010.)

Dès lors, Shadmehr et al. (2010) déterminent, pour une amplitude saccadique donnée, la durée optimale du mouvement, telle qu'elle minimise les différents coûts. Les auteurs constatent que leur modèle, appliqué à différentes amplitudes saccadiques, prédit mieux la *main sequence* que celui de Harris et Wolpert (2006) en ce qui concerne les grandes amplitudes.

Shadmehr et ses collègues s'appuient également sur des données expérimentales obtenues dans leur laboratoire pour montrer que la valeur (subjective) des conséquences visuelles des saccades serait intégrée dans les calculs effectués par le système. Ainsi, dans cette étude de 2010, ils opérationnalisent la dévaluation des conséquences en introduisant un délai compris entre 0.1 et 1 seconde entre la fin des saccades et l'affichage de la cible (qu'ils éteignent préalablement au cours du déplacement de l'œil). Les auteurs observent que le pic de vitesse diminue de manière hyperbolique en fonction de ce délai. Selon eux, cela étayerait leur définition du coût de la durée saccadique.

La valeur (subjective) des conséquences saccadiques a été également manipulée en faisant suivre les saccades de la présentation de différentes images. La recherche de Xu-Wilson et al., (2009) montre ainsi que le pic de vitesse augmente avec la valeur de ces images. Ces résultats sont corroborés par la recherche récente de Reppert, Choi, Haith et Shadmehr (2012). Dans trois conditions expérimentales successives, ils font suivre les saccades soit d'une image constituée de points et dépourvue de signification, soit d'une image représentant un seul et même visage humain, soit d'une image représentant aléatoirement un visage parmi plusieurs. D'une part, les pics de vitesse sont corrélés avec la « valeur » (subjective) de ces conséquences ; ils sont les plus élevés lorsque les images varient. D'autre part, après une augmentation transitoire lors de la présentation des images de visages, ces pics de vitesse diminuent ensuite progressivement, si bien qu'à la fin de l'expérience ils se situent en dessous du niveau de base quel que soit le type d'images présentées. Reppert et al. (2012) interprètent ces résultats comme le reflet de la dévaluation des images suite à leurs présentations répétées.

Enfin, Reppert et al. (2012) examinent également l'évolution de la variabilité des gains saccadiques. Constatant une relation inverse entre une augmentation de cette variabilité et la baisse du pic de vitesse pour une amplitude constante, ils en concluent que le compromis entre vitesse et précision, postulé par Harris et Wolpert (1998, 2006) et sous-tendant bon nombre de modèles computationnels, serait en fait flexible, « modifiable de manière continue » (p. 4).

Les données de notre deuxième étude ne contestent pas le fait que les propriétés saccadiques puissent être modulées indépendamment les unes des autres : la variabilité des amplitudes des saccades y a été manipulée sans que cela se soit accompagné d'un changement de gain médian ni de pic de vitesse.

En revanche, comme cela a été exposé dans la première partie de cette thèse, l'analyse du comportement propose une approche différente de l'origine des changements, qui réside selon elle dans la modification du contrôle exercé par l'environnement, et non dans d'éventuels calculs internes (Donahoe, 2003; Skinner, 1953, 1981).

Nonobstant la supputation de modèles et d'opérations internes, l'idée que la durée des mouvements puisse avoir un coût ne semble pas incompatible avec une approche sélectionniste du contrôle des saccades. En effet – qu'il s'agisse d'une contrainte en terme de temps pendant lequel la vision est dégradée (Harris, 1995; Harris & Wolpert, 1998, 2006) ou en terme de dévaluation des conséquences des saccades (Shadmehr et al., 2010) – le système oculo-moteur pourrait éviter les saccades trop lentes : si le temps passé à réaliser les mouvements fonctionne comme une conséquence impliquée dans un processus de renforcement négatif.

Rappelons que ce dernier est défini comme une augmentation de la probabilité des réponses suivies du retrait d'une conséquence aversive. Comme dans le renforcement positif, la valeur (renforçante) de la conséquence est déterminée en fonction de ses effets sur le comportement.

Pour l'instant, cette fonction de la durée des saccades n'est donc qu'une conjecture. Il se pourrait que cet élément ne joue aucun rôle dans le contrôle des saccades, contrairement aux conséquences visuelles qui seraient les seules à être déterminantes. Pour tester l'hypothèse de ce processus de renforcement négatif, nous pourrions éteindre la cible pendant l'émission de la saccade et la réafficher à l'issue d'un délai comme l'ont fait Shadmehr et al. (2010). Mais au lieu d'utiliser un délai constant, il pourrait être proportionnel à la durée des saccades.

Si effectivement cette durée est aversive, ce procédé pourrait faire augmenter progressivement la probabilité des saccades dont la durée est courte, avec comme corollaire éventuel l'augmentation des pics de vitesse. Cela permettrait d'éliminer des facteurs tels que la fatigue ou l'habituation, qui pourraient tout à fait expliquer les diminutions des pics de vitesse observées par Shadmehr et al. (2010) et Reppert et al. (2012) ; en d'autres termes, cela permettrait de tester la *fonction* de cette conséquence.

III. Conclusion

Les recherches présentées dans cette thèse attestent que le contrôle de l'amplitude des saccades peut relever (en partie) d'un mécanisme général d'apprentissage – un apprentissage opérant. En effet, l'application de programmes de renforcement a permis de façonner la distribution des amplitudes saccadiques : leur médiane et leur dispersion ont été manipulées indépendamment l'une de l'autre. Par conséquent, ces travaux se distinguent de ceux des théories computationnelles classiques qui postulent l'implication de mécanismes spécialisés, calibrant des gains internes ou optimisant des fonctions de coûts.

Autrement dit, nos recherches témoignent de la fonction déterminante des conséquences des saccades dans le contrôle de ces dernières. Que ces conséquences soient de nature auditive (comme dans nos expériences) ou visuelle (quand elles correspondent à la vision nette de la cible) ce serait elles qui, indépendamment de tout postulat portant sur des calculs internes, pourraient s'avérer essentielles pour acquérir et maintenir la précision des saccades tout au long de la vie.

REFERENCES

- Abel, L. A., Schmidt, D., Dell'Osso, L. F., & Daroff, R. B. (1978). Saccadic system plasticity in humans. *Ann Neurol*, 4(4), 313-318.
- Abreu-Rodrigues, J., Lattal, K. A., dos Santos, C. V., & Matos, R. A. (2005). Variation, repetition, and choice. *J Exp Anal Behav*, 83(2), 147-168.
- Alahyane, N., & Pélisson, D. (2004). Eye position specificity of saccadic adaptation. *Invest Ophthalmol Vis Sci*, 45(1), 123-130.
- Alahyane, N., & Pélisson, D. (2005). Retention of saccadic adaptation in humans. *Ann N Y Acad Sci*, 1039, 558-562.
- Albano, J. E., & King, W. M. (1989). Rapid adaptation of saccadic amplitude in humans and monkeys. *Invest Ophthalmol Vis Sci*, 30(8), 1883-1893.
- Antonitis, J. J. (1951). Response variability in the white rat during conditioning, extinction, and reconditioning. *J Exp Psychol*, 42(4), 273-281.
- Aslin, R. N., & Salapatek, P. (1975). Saccadic localization of visual targets by the very young human infant. *Attention, Perception & Psychophysics*, 17(3), 293-302.
- Bachelard, G. (1938). *La formation de l'esprit scientifique*. Paris: Vrin.
- Baer, D. M., & Rosales-Ruiz, J. (2003). In the analysis of behavior, what does “develop” mean? . In K. A. Lattal & P. N. Chase (Eds.), *Behavior theory and philosophy* (pp. 339-346). New York: Kluwer Academic/Plenum Press.
- Bahcall, D. O., & Kowler, E. (2000). The control of saccadic adaptation: implications for the scanning of natural visual scenes. *Vision Res*, 40(20), 2779-2796.
- Bahill, A. T., Clark, M. R., & Stark, L. (1975). The main sequence, a tool for studying eye movements. *Mathematical Biosciences*, 24, 181-204.
- Bayer, H. M., & Glimcher, P. W. (2005). Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron*, 47(1), 129-141.
- Becker, W. (1989). Metrics. In R. H. Wurtz & M. E. Goldberg (Eds.), *The neurobiology of saccadic eye movements* (Vol. 3, pp. 13-67).
- Berger, R. J. (1968). Operant conditioning of eye movement in the monkey (*Macaca nemestrina*). *J Exp Anal Behav*, 11(3), 311-320.
- Berlyne, D. E. (1972). Reinforcement values of visual patterns compared through concurrent performances. *J Exp Anal Behav*, 18(2), 281-285.

- Bizo, L. A., & Doolan, K. E. (2008). *Reinforced Behavioral Variability in Humans*. Poster présenté à la 34^{ème} Convention Annuelle de l'ABAI, Chicago, IL, USA.
- Blough, D. S. (1966). The reinforcement of least-frequent interresponse times. *J Exp Anal Behav*, 9(5), 581-591.
- Bollen, E., Bax, J., van Dijk, J. G., Koning, M., Bos, J. E., Kramer, C. G., et al. (1993). Variability of the main sequence. *Invest Ophthalmol Vis Sci*, 34(13), 3700-3704.
- Bonnetblanc, F., & Baraduc, P. (2007). Saccadic adaptation without retinal postsaccadic error. *Neuroreport*, 18(13), 1399-1402.
- Boulanger, B., Ingebos, A. M., Lahak, M., Machado, A., & Richelle, M. (1987). Variabilité comportementale et conditionnement opérant chez l'animal. *L'année psychologique*, 87(3), 417-434.
- Chen-Harris, H., Joiner, W. M., Ethier, V., Zee, D. S., & Shadmehr, R. (2008). Adaptive control of saccades via internal feedback. *J Neurosci*, 28(11), 2804-2813.
- Cherot, C., Jones, A., & Neuringer, A. (1996). Reinforced variability decreases with approach to reinforcers. *J Exp Psychol Anim Behav Process*, 22(4), 497-508.
- Collins, T., & Doré-Mazars, K. (2009). La plasticité de la transformation sensori-motrice dans le système visuel: l'adaptation saccadique. *L'année psychologique*, 109, 509-549.
- Collins, T., Semroud, A., Orriols, E., & Doré-Mazars, K. (2008). Saccade dynamics before, during, and after saccadic adaptation in humans. *Invest Ophthalmol Vis Sci*, 49(2), 604-612.
- Collins, T., Vergilino-Perez, D., Beauvillain, C., & Doré-Mazars, K. (2007). Saccadic adaptation depends on object selection: evidence from between- and within-object saccadic eye movements. *Brain Res*, 1152, 95-105.
- Collins, T., & Wallman, J. (2011). *The role of efference copy in saccadic adaptation*. Paper presented at the Vision Sciences Society.
- Cooper, J. O., Heron, T. E., & Heward, W. L. (2007). Shaping. In *Applied Behavior Analysis* (2nd ed.). Columbus, OH: Pearson Education.
- Darcheville, J. C., Madelain, L., Buquet, C., Charlier, J., & Miossec, Y. (1999). Operant conditioning of the visual smooth pursuit in young infants. *Behav Processes*, 46(2), 131-139.
- Denney, J., & Neuringer, A. (1998). Behavioral variability is controlled by discriminative stimuli. *Animal Learning & Behavior*, 26(2), 154-162.
- Dodge, R. (1903). Five types of eye movement in the horizontal meridian plane of the field of regard. *American Journal of Physiology--Legacy Content*, 8(4), 307-329 %@ 0002-9513.
- Domjan, M. (2009). Instrumental conditioning: foundations. In *The principles of learning and behavior* (6 ed., pp. 143-184). Belmont, CA: Wadsworth.

- Donahoe, J. W. (2003). Selectionism. In K. A. Lattal & P. N. Chase (Eds.), *Behavior theory and philosophy* (pp. 103-128). New York: Kluwer Academic/Plenum Press.
- Donahoe, J. W., Burgos, J. E., & Palmer, D. C. (1993). A selectionist approach to reinforcement. *J Exp Anal Behav*, 60(1), 17-40.
- Donahoe, J. W., & Palmer, D. C. (1994). Selection of behavior. In *Learning and complex behavior* (pp. 31-68). Richmond, MA: Pearson Education.
- Eckstein, M. P., Beutter, B. R., & Stone, L. S. (2001). Quantifying the performance limits of human saccadic targeting during visual search. *Perception*, 30(11), 1389-1401.
- Efron, B., Jolivet, E., & Hordan, R. (1995). *Le bootstrap et ses applications: Discrimination & régression*. Saint-Mandé (France): CISIA.
- Ethier, V., Zee, D. S., & Shadmehr, R. (2008). Changes in control of saccades during gain adaptation. *J Neurosci*, 28(51), 13929-13937.
- Faisal, A. A., Selen, L. P., & Wolpert, D. M. (2008). Noise in the nervous system. *Nat Rev Neurosci*, 9(4), 292-303.
- Francisco, M. T., Borrero, J. C., & Sy, J. R. (2008). Evaluation of absolute and relative reinforcer value using progressive-ratio schedules. *J Appl Behav Anal*, 41(2), 189-202.
- Frens, M. A., & van Opstal, A. J. (1994). Transfer of short-term adaptation in human saccadic eye movements. *Exp Brain Res*, 100(2), 293-306.
- Fuchs, A. F., & Binder, M. D. (1983). Fatigue resistance of human extraocular muscles. *J Neurophysiol*, 49(1), 28-34.
- Fujita, M., Amagai, A., Minakawa, F., & Aoki, M. (2002). Selective and delay adaptation of human saccades. *Brain Res Cogn Brain Res*, 13(1), 41-52.
- Galbicka, G. (1994). Shaping in the 21st century: Moving percentile schedules into applied settings. *J Appl Behav Anal*, 27(4), 739-760.
- Galbicka, G., Kautz, M. A., & Jagers, T. (1993). Response acquisition under targeted percentile schedules: a continuing quandary for molar models of operant behavior. *J Exp Anal Behav*, 60(1), 171-184.
- Galizio, M. (1979). Contingency-shaped and rule-governed behavior: instructional control of human loss avoidance. *J Exp Anal Behav*, 31(1), 53-70.
- Gaveau, V., Martin, O., Prablanc, C., Pélisson, D., Urquizar, C., & Desmurget, M. (2003). On-line modification of saccadic eye movements by retinal signals. *Neuroreport*, 14(6), 875-878.
- Geisler, W. S., Perry, J. S., & Najemnik, J. (2006). Visual search: the role of peripheral information measured using gaze-contingent displays. *J Vis*, 6(9), 858-873.

- Gewirtz, J. L., & Pelaez, M. (1996). In the context of gross environmental and organismic changes, learning provides the main basis for behavioral development. In S. Bijou & E. Ribes (Eds.), *New directions in behavioral development* (pp. 15-34). Reno, NV: Context Press.
- Gharib, A., Gade, C., & Roberts, S. (2004). Control of variation by reward probability. *J Exp Psychol Anim Behav Process*, 30(4), 271-282.
- Glimcher, P. W. (2003). The neurobiology of visual-saccadic decision making. *Annu Rev Neurosci*, 26, 133-179.
- Goetz, E. M., & Baer, D. M. (1973). Social control of form diversity and the emergence of new forms in children's blockbuilding. *J Appl Behav Anal*, 6(2), 209-217.
- Goldberg, M. E., & Wurtz, R. H. (1972). Activity of Superior Colliculus in Behaving Monkey. I. Visual receptive fields of single neurons. *Journal of Neurophysiology*, 35(4), 542-559.
- Grunow, A., & Neuringer, A. (2002). Learning to vary and varying to learn. *Psychon Bull Rev*, 9(2), 250-258.
- Guerra, L. G. G. C., & Silva, M. T. A. (2010). Learning processes and the neural analysis of conditioning. *Psychology & Neuroscience*, 3(2), 195-208.
- Guthrie, B. L., Porter, J. D., & Sparks, D. L. (1983). Corollary discharge provides accurate eye position information to the oculomotor system. *Science*, 221(4616), 1193-1195.
- Harris, C. M. (1995). Does saccadic undershoot minimize saccadic flight-time? A Monte-Carlo study. *Vision Res*, 35(5), 691-701.
- Harris, C. M., & Wolpert, D. M. (1998). Signal-dependent noise determines motor planning. *Nature*, 394(6695), 780-784.
- Harris, C. M., & Wolpert, D. M. (2006). The main sequence of saccades optimizes speed-accuracy trade-off. *Biol Cybern*, 95(1), 21-29.
- Havermann, K., & Lappe, M. (2010). The influence of the consistency of postsaccadic visual errors on saccadic adaptation. *J Neurophysiol*, 103(6), 3302-3310.
- Henson, D. B. (1978). Corrective saccades: effects of altering visual feedback. *Vision Res*, 18(1), 63-67.
- Herman, J. P., Harwood, M. R., & Wallman, J. (2009). Saccade adaptation specific to visual context. *J Neurophysiol*, 101(4), 1713-1721.
- Herrnstein, R. J. (1961). Relative and absolute strength of response as a function of frequency of reinforcement. *J Exp Anal Behav*, 4, 267-272.
- Hikosaka, O., Nakamura, K., & Nakahara, H. (2006). Basal ganglia orient eyes to reward. *J Neurophysiol*, 95(2), 567-584.

- Hodos, W., & Kalman, G. (1963). Effects of increment size and reinforcer volume on progressive ratio performance. *J Exp Anal Behav*, 6, 387-392.
- Hollerman, J. R., Tremblay, L., & Schultz, W. (1998). Influence of reward expectation on behavior-related neuronal activity in primate striatum. *J Neurophysiol*, 80(2), 947-963.
- Hopkinson, J., & Neuringer, A. (2003). Modifying behavioral variability in moderately depressed students. *Behav Modif*, 27(2), 251-264.
- Hopp, J. J., & Fuchs, A. F. (2004). The characteristics and neuronal substrate of saccadic eye movement plasticity. *Prog Neurobiol*, 72(1), 27-53.
- Hopp, J. J., & Fuchs, A. F. (2010). Identifying sites of saccade amplitude plasticity in humans: transfer of adaptation between different types of saccade. *Exp Brain Res*, 202(1), 129-145.
- Ikeda, T., & Hikosaka, O. (2003). Reward-dependent gain and bias of visual responses in primate superior colliculus. *Neuron*, 39(4), 693-700.
- Ingvaldsen, R. P., & Whiting, H. T. A. (1997). Modern views on motor skill learning are not 'representative'! *Human Movement Science*, 16, 705-732.
- Irving, E. L., Steinbach, M. J., Lillakas, L., Babu, R. J., & Hutchings, N. (2006). Horizontal saccade dynamics across the human life span. *Invest Ophthalmol Vis Sci*, 47(6), 2478-2483.
- Joyce, J. H., & Chase, P. N. (1990). Effects of response variability on the sensitivity of rule-governed behavior. *J Exp Anal Behav*, 54(3), 251-262.
- Karniel, A. (2011). Open questions in computational motor control. *J Integr Neurosci*, 10(3), 385-411.
- Killeen, P. R., Hanson, S. J., & Osborne, S. R. (1978). Arousal: its genesis and manifestation as response rate. *Psychol Rev*, 85(6), 571-581.
- Kojima, Y., Iwamoto, Y., & Yoshida, K. (2004). Memory of learning facilitates saccadic adaptation in the monkey. *J Neurosci*, 24(34), 7531-7539.
- Kommerell, G., Olivier, D., & Theopold, H. (1976). Adaptive programming of phasic and tonic components in saccadic eye movements. Investigations of patients with abducens palsy. *Invest Ophthalmol*, 15(8), 657-660.
- Körding, K. (2007). Decision theory: what "should" the nervous system do? *Science*, 318(5850), 606-610.
- Körding, K., Tenenbaum, J. B., & Shadmehr, R. (2007). The dynamics of memory as a consequence of optimal adaptation to a changing body. *Nat Neurosci*, 10(6), 779-786.
- Krauzlis, R. (2008). Eye movements. In L. R. Squire & D. Berg (Eds.), *Fundamental Neuroscience* (3rd ed., pp. 775-792). Amsterdam: Elsevier.

- Land, M. F. (1999). Motion and vision: why animals move their eyes. *J Comp Physiol A*, 185(4), 341-352.
- Lattal, K. A., & Neef, N. A. (1996). Recent reinforcement-schedule research and applied behavior analysis. *J Appl Behav Anal*, 29(2), 213-230.
- Lauwereyns, J., Watanabe, K., Coe, B., & Hikosaka, O. (2002). A neural correlate of response bias in monkey caudate nucleus. *Nature*, 418(6896), 413-417.
- Lebedev, S., Van Gelder, P., & Tsui, W. H. (1996). Square-root relations between main saccadic parameters. *Invest Ophthalmol Vis Sci*, 37(13), 2750-2758.
- Lee, R., McComas, J. J., & Jawor, J. (2002). The effects of differential and lag reinforcement schedules on varied verbal responding by individuals with autism. *J Appl Behav Anal*, 35(4), 391-402.
- Lee, R., Sturmey, P., & Fields, L. (2007). Schedule-induced and operant mechanisms that influence response variability: A review and implications for future investigations. *The Psychological Report*, 57(3), 429-455.
- Leigh, R. J., & Kennard, C. (2004). Using saccades as a research tool in the clinical neurosciences. *Brain*, 127(Pt 3), 460-477.
- Lewis, R. F., Zee, D. S., Hayman, M. R., & Tamargo, R. J. (2001). Oculomotor function in the rhesus monkey after deafferentation of the extraocular muscles. *Exp Brain Res*, 141(3), 349-358.
- Li, X., & Basso, M. A. (2008). Preparing to move increases the sensitivity of superior colliculus neurons. *The Journal of Neuroscience*, 28(17), 4561-4577.
- Machado, A. (1989). Operant conditioning of behavioral variability using a percentile reinforcement schedule. *J Exp Anal Behav*, 52(2), 155-166.
- Machado, A. (1992). Behavioral variability and frequency-dependent selection. *J Exp Anal Behav*, 58(2), 241-263.
- Madelain, L., Champrenaut, L., & Chauvin, A. (2007). Control of sensorimotor variability by consequences. *J Neurophysiol*, 98(4), 2255-2265.
- Madelain, L., Harwood, M. R., Herman, J. P., & Wallman, J. (2010). Saccade adaptation is unhampered by distractors. *J Vis*, 10(12), 29.
- Madelain, L., & Krauzlis, R. J. (2003). Effects of learning on smooth pursuit during transient disappearance of a visual target. *J Neurophysiol*, 90(2), 972-982.
- Madelain, L., & Paeye, C. (2011). *Changes in saccadic peak velocity induced by reinforcement learning*. Poster présenté à l'ECCEM 2011, Marseille, France.

- Madelain, L., Paeye, C., & Darcheville, J. C. (2011). Operant control of human eye movements. *Behav Processes*, 87(1), 142-148.
- Madelain, L., Paeye, C., & Wallman, J. (2011). Modification of saccadic gain by reinforcement. *J Neurophysiol*, 106(1), 219-232.
- Maes, J. H. (2003). Response stability and variability induced in humans by different feedback contingencies. *Learn Behav*, 31(4), 332-348.
- Maes, J. H., & Van der Goot, M. (2006). Human operant learning under concurrent reinforcement of response variability. *Learning and Motivation*, 37(1), 79-92.
- McLaughlin. (1967). Parametric adjustment in saccadic eye movements. *Perception & Psychophysics*, 2, 359-362.
- Miller, J. M., Anstis, T., & Templeton, W. B. (1981). Saccadic plasticity: parametric adaptive control by retinal feedback. *J Exp Psychol Hum Percept Perform*, 7(2), 356-366.
- Miller, N., & Neuringer, A. (2000). Reinforcing variability in adolescents with autism. *J Appl Behav Anal*, 33(2), 151-165.
- Montagnini, A., & Chelazzi, L. (2005). The urgency to look: prompt saccades to the benefit of perception. *Vision Res*, 45(27), 3391-3401.
- Montague, P. R., Hyman, S. E., & Cohen, J. D. (2004). Computational roles for dopamine in behavioural control. *Nature*, 431(7010), 760-767.
- Morvan, C., & Maloney, L. T. (2012). Human visual search does not maximize the post-saccadic probability of identifying targets. *PLoS One*, 8(2), e1001342.
- Myers, C. W., & Gray, W. D. (2010). Visual scan adaptation during repeated visual search. *J Vis*, 10(8), 4.
- Myerson, J., & Green, L. (1995). Discounting of delayed rewards: Models of individual choice. *J Exp Anal Behav*, 64(3), 263-276.
- Najemnik, J., & Geisler, W. S. (2005). Optimal eye movement strategies in visual search. *Nature*, 434(7031), 387-391.
- Najemnik, J., & Geisler, W. S. (2008). Eye movement statistics in humans are consistent with an optimal search strategy. *J Vis*, 8(3), 4 1-14.
- Nakamura, K., & Hikosaka, O. (2006). Role of dopamine in the primate caudate nucleus in reward modulation of saccades. *J Neurosci*, 26(20), 5360-5369.
- Navalpakkam, V., Koch, C., Rangel, A., & Perona, P. (2010). Optimal reward harvesting in complex perceptual environments. *Proc Natl Acad Sci U S A*, 107(11), 5232-5237.
- Neuringer, A. (1986). Can people behave "randomly?": The role of feedback. *Journal of Experimental Psychology: General*, 115(1), 62-75.

- Neuringer, A. (1992). Choosing to vary and repeat. *Psychological Science*, 3(4), 246-250.
- Neuringer, A. (1993). Reinforced variation and selection. *Animal Learning & Behavior*, 21(2), 83-91.
- Neuringer, A. (2002). Operant variability: evidence, functions, and theory. *Psychon Bull Rev*, 9(4), 672-705.
- Neuringer, A. (2009). Operant Variability and the Power of Reinforcement. *The Behavior Analyst Today*, 10(2), 319-342.
- Neuringer, A., Deiss, C., & Olson, G. (2000). Reinforced variability and operant learning. *J Exp Psychol Anim Behav Process*, 26(1), 98-111.
- Neuringer, A., & Jensen, G. (2010). Operant variability and voluntary action. *Psychol Rev*, 117(3), 972-993.
- Neuringer, A., Kornell, N., & Olufs, M. (2001). Stability and variability in extinction. *J Exp Psychol Anim Behav Process*, 27(1), 79-94.
- Neuringer, A., & Voss, C. (1993). Approximating chaotic behavior. *Psychological Science*, 4(2), 113-119.
- Noto, C. T., & Robinson, F. R. (2001). Visual error is the stimulus for saccade gain adaptation. *Brain Res Cogn Brain Res*, 12(2), 301-305.
- Noto, C. T., & Robinson, F. R. (2005). Saccades to targets in a large visual scene do not adapt in monkey. Poster présenté à la Society for Neuroscience, Washington, DC, USA.
- Noto, C. T., Watanabe, S., & Fuchs, A. F. (1999). Characteristics of simian adaptation fields produced by behavioral changes in saccade size and direction. *J Neurophysiol*, 81(6), 2798-2813.
- Paeye, C., Darcheville, J. C., & Madelain, L. (2007). *Shaping induces long-term enhancements in pursuit-like-smooth movements under image-stabilization condition*. Poster présenté à l'ECVP 2007, Arezzo, Italy. *ECVP - Perception*, 36, *ECVP Abstract Supplement*.
- Page, S., & Neuringer, A. (1985). Variability is an operant. *Journal of Experimental Psychology: Animal Behavior Processes*, 11(3), 429-452.
- Panouillères, M., Urquizar, C., Salemme, R., & Pélisson, D. (2011). Sensory processing of motor inaccuracy depends on previously performed movement and on subsequent motor corrections: a study of the saccadic system. *PLoS One*, 6(2), e17329.
- Pelaez, M., Gewirtz, J. L., & Wong, S. E. (2008). A critique of stage theories of human development. In B. A. Thyer, K. M. Sowers & C. N. Dulmus (Eds.), *Comprehensive handbook of social work and social welfare* (pp. 503-518). Hoboken, New Jersey: Wiley & Sons.
- Pélisson, D., Alahyane, N., Panouillères, M., & Tilikete, C. (2010). Sensorimotor adaptation of saccadic eye movements. *Neurosci Biobehav Rev*, 34(8), 1103-1120.

- Platt, M. L., & Glimcher, P. W. (1999). Neural correlates of decision variables in parietal cortex. *Nature*, 400(6741), 233-238.
- Poling, A., Edwards, T. L., Weeden, M., & Foster, T. M. (2011). The matching law. *The Psychological Record*(61), 313-322.
- Pryor, K. W., Haag, R., & O'Reilly, J. (1969). The creative porpoise: training for novel behavior. *J Exp Anal Behav*, 12(4), 653-661.
- Reddi, B. A., & Carpenter, R. H. (2000). The influence of urgency on decision time. *Nat Neurosci*, 3(8), 827-830.
- Reppert, T. R., Choi, J. E., Haith, A. M., & Shadmehr, R. (2012). *Changes in saccade kinematics associated with the value and novelty of a stimulus*. Communication présentée à la Conference on Information Sciences and Systems.
- Rescorla, R., & Wagner, A. (1972). A Theory of Pavlovian Conditioning: Variations in the Effectiveness of Reinforcement and Nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: current research and theory* (pp. 64-99).
- Robinson, F. R., Noto, C. T., & Bevans, S. E. (2003). Effect of visual error size on saccade adaptation in monkey. *J Neurophysiol*, 90(2), 1235-1244.
- Robinson, F. R., Soetedjo, R., & Noto, C. (2006). Distinct short-term and long-term adaptation to reduce saccade size in monkey. *J Neurophysiol*, 96(3), 1030-1041.
- Rosenberger, P. B. (1973). Concurrent schedule control of human visual target fixations. *J Exp Anal Behav*, 20(3), 411-416.
- Ross, C., & Neuringer, A. (2002). Reinforcement of variations and repetitions along three independent response dimensions. 57, 199-209.
- Sato, M., & Hikosaka, O. (2002). Role of primate substantia nigra pars reticulata in reward-oriented saccadic eye movement. *J Neurosci*, 22(6), 2363-2373.
- Schroeder, S. R., & Holland, J. G. (1968). Operant control of eye movements. *J Appl Behav Anal*, 1(2), 161-166.
- Schroeder, S. R., & Holland, J. G. (1969). Reinforcement of eye movement with concurrent schedules. *J Exp Anal Behav*, 12(6), 897-903.
- Schultz, W. (2000). Multiple reward signals in the brain. *Nat Rev Neurosci*, 1(3), 199-207.
- Schultz, W. (2006). Behavioral theories and the neurophysiology of reward. *Annu Rev Psychol*, 57, 87-115.
- Schultz, W. (2010). Dopamine signals for reward value and risk: basic and recent data. *Behav Brain Funct*, 6, 24.

- Schütz, A. C., Braun, D. I., & Gegenfurtner, K. R. (2011). Eye movements and perception: a selective review. *J Vis*, 11(5).
- Schütz, A. C., & Gegenfurtner, K. R. (2010). *Dynamic integration of saliency and reward information for saccadic eye movements*. Poster présenté à VSS 2010, Naples, FL, USA. *J Vis*, 10(7).
- Scudder, C. A., Batourina, E. Y., & Tunder, G. S. (1998). Comparison of two methods of producing adaptation of saccade size and implications for the site of plasticity. *J Neurophysiol*, 79(2), 704-715.
- Seeberger, T., Noto, C., & Robinson, F. (2002). Non-visual information does not drive saccade gain adaptation in monkeys. *Brain Res*, 956(2), 374-379.
- Semmlow, J. L., Gauthier, G. M., & Vercher, J. L. (1989). Mechanisms of short-term saccadic adaptation. *J Exp Psychol Hum Percept Perform*, 15(2), 249-258.
- Shadmehr, R., Orban de Xivry, J. J., Xu-Wilson, M., & Shih, T. Y. (2010). Temporal discounting of reward and the cost of time in motor control. *J Neurosci*, 30(31), 10507-10516.
- Shelhamer, M., & Clendaniel, R. (2002). Sensory, motor, and combined contexts for context-specific adaptation of saccade gain in humans. *Neurosci Lett*, 332(3), 200-204.
- Skinner, B. F. (1953). *Science and human behavior*. New-York: Macmillan.
- Skinner, B. F. (1981). Selection by consequences. *Science*, 213(4507), 501-504.
- Sohn, J. W., & Lee, D. (2006). Effects of reward expectancy on sequential eye movements in monkeys. *Neural Netw*, 19(8), 1181-1191.
- Souza, A. S., Pontes, T. N., & Abreu-Rodrigues, J. (2012). Varied but not necessarily random: Human performance under variability contingencies is affected by instructions. *Learn Behav*.
- Stahlman, W. D., Roberts, S., & Blaisdell, A. P. (2010). Effect of reward probability on spatial and temporal variation. *J Exp Psychol Anim Behav Process*, 36(1), 77-91.
- Stokes, P. D. (1995). Learned variability. *Animal Learning & Behavior*, 23(2), 164-176.
- Stokes, P. D., & Balsam, P. (2001). An optimal period for setting sustained variability levels. *Psychon Bull Rev*, 8(1), 177-184.
- Straube, A., Fuchs, A. F., Usher, S., & Robinson, F. R. (1997). Characteristics of saccadic gain adaptation in rhesus macaques. *J Neurophysiol*, 77(2), 874-895.
- Stritzke, M., Trommershäuser, J., & Gegenfurtner, K. R. (2009). Effects of salience and reward information during saccadic decisions under risk. *J Opt Soc Am A Opt Image Sci Vis*, 26(11), B1-13.
- Sugrue, L. P., Corrado, G. S., & Newsome, W. T. (2004). Matching behavior and the representation of value in the parietal cortex. *Science*, 304(5678), 1782-1787.

- Takikawa, Y., Kawagoe, R., Itoh, H., Nakahara, H., & Hikosaka, O. (2002). Modulation of saccadic eye movements by predicted reward outcome. *Exp Brain Res*, 142(2), 284-291.
- Tatler, B. W., Hayhoe, M. M., Land, M. F., & Ballard, D. H. (2011). Eye guidance in natural vision: reinterpreting salience. *J Vis*, 11(5), 5.
- Tian, J., & Zee, D. S. (2010). Context-specific saccadic adaptation in monkeys. *Vision Res*, 50(23), 2403-2410.
- Trommershäuser, J., Glimcher, P. W., & Gegenfurtner, K. R. (2009). Visual processing, learning and feedback in the primate eye movement system. *Trends Neurosci*, 32(11), 583-590.
- van Beers, R. J. (2007). The sources of variability in saccadic eye movements. *J Neurosci*, 27(33), 8757-8770.
- van Beers, R. J. (2008). Saccadic eye movements minimize the consequences of motor noise. *PLoS One*, 3(4), e2070.
- Wallman, J., & Fuchs, A. F. (1998). Saccadic gain modification: visual error drives motor adaptation. *J Neurophysiol*, 80(5), 2405-2416.
- Watanabe, K., Lauwereyns, J., & Hikosaka, O. (2003). Neural correlates of rewarded and unrewarded eye movements in the primate caudate nucleus. *J Neurosci*, 23(31), 10052-10057.
- Wolpert, D. M., & Ghahramani, Z. (2000). Computational principles of movement neuroscience. *Nat Neurosci*, 3 Suppl, 1212-1217.
- Wolpert, D. M., Ghahramani, Z., & Flanagan, J. R. (2001). Perspectives and problems in motor learning. *Trends Cogn Sci*, 5(11), 487-494.
- Wolpert, D. M., Ghahramani, Z., & Jordan, M. I. (1995). An internal model for sensorimotor integration. *Science*, 269(5232), 1880-1882.
- Wolpert, D. M., & Kawato, M. (1998). Multiple paired forward and inverse models for motor control. *Neural Netw*, 11(7-8), 1317-1329.
- Wong, A. L., & Shelhamer, M. (2011). Sensorimotor adaptation error signals are derived from realistic predictions of movement outcomes. *J Neurophysiol*, 105(3), 1130-1140.
- Xu-Wilson, M., Zee, D. S., & Shadmehr, R. (2009). The intrinsic value of visual information affects saccade velocities. *Exp Brain Res*, 196(4), 475-481.
- Yuodelis, C., & Hendrickson, A. (1986). A qualitative and quantitative analysis of the human fovea during development. *Vision Res*, 26(6), 847-855.

ANNEXES

Annexe 1 : Formulaire de consentement

Je, soussigné(e) :

Nom : Prénom :

Adresse :

Code Postal.....Ville :

Tel :Age :ans

ayant déjà participé à une expérience sur les mouvements oculaires : ☐ Oui ☐ Non

portant des lunettes ou lentilles : ☐ Oui ☐ Non

déclare accepter, librement et de façon éclairée, de participer à l'étude portant sur le *contrôle des mouvements oculaires par les contingences environnementales*.

L'étudiante Céline PAEYE, en Doctorat de psychologie, m'a proposé de participer à une recherche qui a pour but d'étudier l'apprentissage d'un comportement oculaire grâce à une technique comportementale.

Responsable scientifique : Laurent MADELAIN.

A cette fin, il m'est demandé de regarder une cible apparaissant transitoirement sur un écran. Les mouvements oculaires seront mesurés par une caméra à infrarouges. Pendant l'expérience, la tête est stabilisée par un système de maintien de la mâchoire supérieure.

L'étude se déroule sur une à deux semaines, à raison de sessions quotidiennes d'environ minutes (séparées par une pause d'une dizaine de minutes).

- Je suis libre d'accepter ou de refuser de participer à cette étude.
- Je peux retirer à tout moment ma participation à cette recherche, cela quelles que soient mes raisons et sans que cela ait quelque conséquence négative que ce soit.
- Les données me concernant seront utilisées à des fins scientifiques et resteront strictement anonymes et confidentielles.
- J'ai la possibilité d'obtenir des informations supplémentaires concernant cette étude auprès de l'investigateur principal, et ce dans la limite des contraintes du plan de recherche.
- Mon consentement ne décharge en rien l'investigateur de ses responsabilités morales et légales et je conserve mes droits garantis par la loi.

Fait à, le, en 2 exemplaires.

Signature de l'investigateur :

Signature du participant :

Annexe 2 : Consignes de l'étude 1

a. Quelle que soit l'expérience (ligne de base)

Tu es en compétition avec les autres participants. Il faut que tu marques le plus de points pour gagner ; le vainqueur recevra un lot personnalisé. Je t'expliquerai plus tard comment gagner des points. On va faire d'abord quelques mesures pour la précision du signal que je vais enregistrer. Tu vas appuyer ta mâchoire supérieure sur l'empreinte pour que ta tête reste totalement immobile pendant toute la durée de l'expérience. Si ta tête bougeait, cela fausserait les données en faisant sortir ton œil de la zone d'enregistrement.

Tu vas devoir regarder le petit point noir au centre du carré blanc. Il va se déplacer à plusieurs endroits. À chaque fois, appuie sur le bouton dès que tu fixeras parfaitement le petit point central. Quand la mesure sera précise, il ne faudra plus bouger, l'expérience commencera.

Tu verras apparaître sur la gauche une croix blanche qui annonce le début d'un essai. Tu vas fixer cette croix jusqu'à ce qu'elle disparaisse. A ce moment-là, une cible va apparaître sur la droite. Tu dois simplement la regarder. Puis un autre essai recommence : tu regardes à nouveau la croix, puis la cible, et ainsi de suite.

b. Adaptations

Expériences 1, 2, 4 et 5 : agent renforçateur auditif

C'est maintenant que le défi commence vraiment. Plus tu entends de « bips », plus tu marqueras de points. Chaque « bravo » signifie que tu as marqué un point. Le but est bien sûr de marquer le plus de points possibles, sachant que tu n'es pas le/la seul(e) participant(e), et que vous avez tous le même but : finir en tête du classement ! Tu n'as pas de question ? On fait les mesures préliminaires et c'est parti !

Expérience 3 : agent renforçateur visuel

C'est maintenant que le défi commence vraiment. Plus tu vois la cible longtemps, plus tu marqueras de points. Chaque « bravo » signifie que tu as marqué un point. Le but est bien sûr de marquer le plus de points possibles, sachant que tu n'es pas le/la seul(e) participant(e), et que vous avez tous le même but : finir en tête du classement ! Tu n'as pas de question ? On fait les mesures préliminaires et c'est parti !

Paradigme conventionnel

Tu verras apparaître sur la gauche une croix blanche qui annonce le début d'un essai. Tu vas fixer cette croix jusqu'à ce qu'elle disparaisse. A ce moment-là, une cible va apparaître sur la droite. Tu dois simplement la regarder. Puis un autre essai recommence : tu regardes à nouveau la croix, puis la cible, et ainsi de suite.

Annexe 3 : Consignes de l'étude 2

Tu es en compétition avec les autres participants. Il faut que tu marques le plus de points pour gagner ; le vainqueur recevra un lot. Je t'expliquerai plus tard comment gagner des points, lors du quatrième bloc. On va faire d'abord quelques mesures pour la précision du signal que je vais enregistrer. Tu vas mordre l'empreinte, et appuyer ton front sur la barre pour que ta tête reste totalement immobile pendant toute la durée de l'expérience. Si ta tête bougeait, cela fausserait les données en faisant sortir ton œil de la zone d'enregistrement.

Tu vas devoir regarder le petit point noir au centre du carré blanc. Il va se déplacer à plusieurs endroits. À chaque fois, appuie sur un des boutons devant toi dès que tu fixeras parfaitement le petit point noir au centre du carré blanc. Quand la mesure sera précise, il ne faudra plus bouger, l'expérience commencera juste après.

a. Avant la ligne de base ordinaire

Tu verras apparaître d'abord, d'un côté de l'écran, une croix blanche qui annonce le début d'un essai. Tu vas fixer cette croix. N'anticipe pas, sinon l'essai s'annule : tu as assez de temps pour regarder la croix jusqu'à ce qu'elle disparaisse. A ce moment-là, une cible, une tâche blanche, va apparaître de l'autre côté de l'écran. Tu dois simplement la regarder. Puis un autre essai recommence : tu regardes à nouveau la croix, puis la cible, et ainsi de suite.

b. Avant la ligne de base stabilisée

C'est maintenant que le défi commence vraiment.

Pour marquer des points, tu dois entendre des « bips ». Après un certain nombre de « bips », tu entendras une voix prononçant une espèce de « bravo » ; chaque « bravo » équivaut à un point. Tu obtiens ces bips – et donc ces « bravos » – en fonction de la variabilité de l'amplitude de tes saccades oculaires. Juste une précision : une saccade correspond à un mouvement rapide des yeux quand tu regardes un point puis un autre. Le but est bien sûr de marquer le plus de points possibles, sachant que tu n'es pas le/la seul(e) participant(e), et que vous avez tous le même but : finir en tête du classement ! Tu n'as pas de question ? On fait les mesures préliminaires et c'est parti !

Annexe 4 : Consignes de l'étude 3

Au début de chaque session, on procède à une calibration. Il faut que vous regardiez les ronds qui vont apparaître à différents endroits. Puis il ne faudra plus bouger car l'expérience commencera.

Il faut bien fixer le centre de l'écran jusqu'à ce que vous entendiez un bip qui marque le début d'un essai.

Des items vont apparaître sur l'écran, des distracteurs et des cibles, qui ont cette forme
(un exemple de configuration d'items est présenté).

Votre tâche est de regarder ces items, et de voir le maximum de cibles.

En même temps que vous verrez une cible, vous entendrez un son, et au bout d'un certain nombre de cibles, vous pouvez gagner des points. Chaque point vous rapportera 10 centimes d'euro, qui seront comptabilisés et versés à la fin de l'expérience.

Il peut y avoir une ou plusieurs cibles par image.

Vous avez des questions ? C'est parti !

Annexe 5 : Articles des études 1 et 2

Madelain, L., Paeye, C., & Wallman, J. (2011). Modification of saccadic gain by reinforcement. *Journal of Neurophysiology*, 106(1), 219-232.

Paeye, C., & Madelain, L. (2011). Reinforcing saccadic amplitude variability. *Journal of the Experimental Analysis of Behavior*, 95(2), 149-162.

Modification of saccadic gain by reinforcement

Laurent Madelain, Céline Paeye and Josh Wallman

J Neurophysiol 106:219-232, 2011. First published 27 April 2011; doi:10.1152/jn.01094.2009

You might find this additional info useful...

Supplemental material for this article can be found at:

<http://jn.physiology.org/content/suppl/2011/04/28/jn.01094.2009.DC1.html>

This article cites 57 articles, 25 of which can be accessed free at:

<http://jn.physiology.org/content/106/1/219.full.html#ref-list-1>

Updated information and services including high resolution figures, can be found at:

<http://jn.physiology.org/content/106/1/219.full.html>

Additional material and information about *Journal of Neurophysiology* can be found at:

<http://www.the-aps.org/publications/jn>

This information is current as of July 18, 2011.

Modification of saccadic gain by reinforcement

Laurent Madelain,¹ Céline Paeye,¹ and Josh Wallman²

¹Laboratoire Ureca, Unité de Formation et de Recherche de Psychologie, Université Ch. De Gaulle Lille III, Villeneuve d'Ascq, France; and ²Department of Biology, City College, City University of New York, New York, New York

Submitted 14 December 2009; accepted in final form 25 April 2011

Madelain L, Paeye C, Wallman J. Modification of saccadic gain by reinforcement. *J Neurophysiol* 106: 219–232, 2011. First published April 27, 2011; doi:10.1152/jn.01094.2009.—Control of saccadic gain is often viewed as a simple compensatory process in which gain is adjusted over many trials by the postsaccadic retinal error, thereby maintaining saccadic accuracy. Here, we propose that gain might also be changed by a reinforcement process not requiring a visual error. To test this hypothesis, we used experimental paradigms in which retinal error was removed by extinguishing the target at the start of each saccade and either an auditory tone or the vision of the target on the fovea was provided as reinforcement after those saccades that met an amplitude criterion. These reinforcement procedures caused a progressive change in saccade amplitude in nearly all subjects, although the rate of adaptation differed greatly among subjects. When we reversed the contingencies and reinforced those saccades landing closer to the original target location, saccade gain changed back toward normal gain in most subjects. When subjects had saccades adapted first by reinforcement and a week later by conventional intrasaccadic step adaptation, both paradigms yielded similar degrees of gain changes and similar transfer to new amplitudes and to new starting positions of the target step as well as comparable rates of recovery. We interpret these changes in saccadic gain in the absence of postsaccadic retinal error as showing that saccade adaptation is not controlled by a single error signal. More generally, our findings suggest that normal saccade adaptation might involve general learning mechanisms rather than only specialized mechanisms for motor calibration.

saccade; adaptation; reinforcement learning; eye movements; plasticity

PRIMATES USE TWO TYPES of voluntary eye movements to view their visual environment: saccades, rapid movements that orient the retinal image of a target toward the high-acuity area of the retina, and pursuit, continuous responses that smoothly track target motion. There is evidence that pursuit might be controlled by learned contingencies (Darcheville et al. 1999; Madelain and Krauzlis 2003), and, therefore, pursuit may be an operant behavior, that is, a response selected and maintained by its consequences (Skinner 1981). Saccades, as well, have attributes of operant behaviors. Saccades made by monkeys to rewarded locations had shorter latencies (Ikeda and Hikosaka 2003; Lauwereyns et al. 2002; Watanabe et al. 2003a, 2003b) and higher peak velocities (Takikawa et al. 2002) than saccades of the same amplitude to nonrewarded locations. In humans, it has recently been shown that saccadic peak velocities increased when the value of the visual stimulus was high (Xu-Wilson et al. 2009) and that saccadic end point variability is sensitive to reward (Paeye and Madelain 2011). This mod-

ulation might be the result of activation of dopaminergic neurons in the basal ganglia (Nakamura and Hikosaka 2006). Specifically, different neurons in the caudate nucleus fire in anticipation of rewarded versus nonrewarded saccades (Watanabe et al. 2003b), midbrain dopaminergic neurons fire in anticipation of rewarded saccades (Bayer and Glimcher 2005), and the basal ganglia modulate the superior colliculus, facilitating saccades to rewarded locations (Hikosaka et al. 2006). Furthermore, if monkeys are rewarded for making saccades to either of two targets, with saccades to each being reinforced on different variable-interval schedules of reinforcement, monkeys learn to “match” their frequency of saccades to each target to its relative frequency of reinforcement (Sugrue et al. 2004).

Because there are costs associated with making saccades, in that vision is impaired during the saccade and the target may be lost among similar visual objects, it is not surprising that the reaction time can be shaped by circumstances. For example, Montagnini and Chelazzi (2005) showed that saccade latencies in humans are reduced when subjects perform a discrimination of the saccade target displayed for a brief interval beginning at the median saccade landing time. (In this situation, the perceptual task was easier when saccade latencies were short.) One interpretation of these results is that the successful performance of the task reinforces the short latencies. Furthermore, saccades in response to a small target step have a longer latency when the subject attending to a large target than to a small target, presumably because less benefit (additional information) is obtained by making the saccade (Harwood et al. 2008; Madelain et al. 2005). More directly, the median and dispersion of saccade latencies can be independently altered by reinforcement (Madelain et al. 2007). These studies suggest that the time of triggering of a saccade can be under operant control. Perhaps under natural circumstances a high-acuity view of the target acts as reinforcement for the saccadic system, thereby shaping subsequent saccades. In the present work, we explored whether the amplitude of saccades is sensitive to reinforcement as well.

Because saccadic eye movements are too rapid to allow online visual feedback control, saccade accuracy is maintained by long-term adjustments of saccade amplitude relative to the distance of the target from the fovea (that is, changes in saccade gain). In the laboratory, the ability to alter saccadic amplitude is generally studied by having a target step away from the fovea and, while the eye is in flight towards the new location, moving the target (for example, slightly back toward its original location) so that the eye does not land on the target but must make a second, corrective, saccade to acquire it. After a number of repetitions of this intrasaccadic step paradigm, the amplitude of the initial saccade decreases, bringing the eye closer to the displaced target location (McLaughlin 1967). This

Address for reprint requests and other correspondence: L. Madelain, Université Ch. De Gaulle Lille III, Villeneuve d'Ascq 59650, France (e-mail: laurent.madelain@univ-lille3.fr).

behavioral plasticity presumably maintains the accuracy of saccades despite changes in the eyes, muscles, and oculomotor system during development and aging or after injuries (Abel et al. 1978; Kommerell et al. 1976; Scudder et al. 1998). Saccadic adaptation reliably changes the amplitude or direction of saccades of various types (targeting saccades, memory-guided saccades, and scanning saccades), albeit with considerable variation from subject to subject and from day to day (for reviews, see Hopp and Fuchs 2004; Pelisson et al. 2010).

Conventionally, saccadic adaptation has been viewed as controlled by a visual position error signal, the "retinal error," i.e., the postsaccadic distance between the target and fovea, which is progressively reduced over many trials (Noto and Robinson 2001; Wallman and Fuchs 1998). Recent studies have shown that the relevant retinal error is that associated with the target rather than the retinal error arising from similar stimuli near the fovea (Madelain et al. 2010). Computational models of sensorimotor control have also postulated a mechanism using retinal error to account for motor learning: a forward model is used to predict the outcome of the movement, and this predicted signal is then compared with the actual sensory outcome of a movement (Kording et al. 2007; Wolpert et al. 1995), with the result of this comparison being used to adjust the motor command. In a similar vein, a visual comparison model has been proposed in which adaptation depends on a comparison of the retinal image after the saccade lands with the image that would be predicted based on the size and direction of the planned saccade (Bahcall and Kowler 2000). A critical feature of these hypotheses is that they imply that adaptation must involve a visual signal. Adaptation is viewed as a machine-like servomechanism in which the position of the target on the fovea is fed back and compared with a predicted or desired signal, with the outcome of this comparison driving the adjustments of the internal gain. Adaptation would therefore constitute a simple calibration process that maintains an accurate mapping between target eccentricity and saccadic motor output.

Here, we propose an alternative view in which a reinforcement process increases the frequency of saccades landing closer to the target and decreases the frequency of those landing farther away from it. To put the matter in the terms of operant conditioning, we suggest that saccadic gain may depend on the cumulative effect of selection by reinforcement (Donahoe et al. 1993). Postulating that a clear view of the target is reinforcing, we would expect that in a conventional situation saccades that land close to the target would be reinforced, whereas those landing far from it would not be. Here, the position error signal is correlated with the reinforcing outcome of the movement. In other words, adaptation would not depend solely on visual signals but could be controlled by an outcome of another nature.

Because both reinforcement and retinal error provide feedback to the oculomotor system about the efficacy of each saccade, one could argue that the difference between our hypothesis and the conventional view is less than it seems. However, an essential difference is that retinal error signals the direction and magnitude of the error on each saccade, whereas reinforcement signals neither. Therefore, if reinforcement can provide adequate feedback to guide saccade adaptation, it would show that these visual signals, fundamental to servomechanisms, are not essential. The goal of this report is to

probe whether changes in saccade amplitudes might be observed using reinforcement procedures, in the absence of postsaccadic retinal errors.

To test for this possibility, we conducted three series of experiments. In *experiments 1* and *2*, the retinal error after the saccades was absent, but the subject was signaled (using an auditory stimulus) when saccades of appropriate size had been executed. Because in these experiments the target was blanked at saccade onset, the visual consequences were disentangled from the appetitive consequences of the saccade. Because we found that reinforcement could induce decreases (*experiment 1*) and increases (*experiment 2*) in saccade amplitude, we further studied the characteristics of reinforcement adaptation. We tested whether signaling saccades of appropriate size by a view of the target in the absence of a retinal error might also induce a change in amplitude (*experiment 3*). Finally, we asked whether the auditory reinforcement would be effective despite the presence of the target image on the fovea (*experiments 4* and *5*). This paradigm was also used to compare some specificities of conventional and reinforcement adaptation.

METHODS

Reinforcement Procedure

To induce modifications of saccade amplitude, we had the target being fixated step 14.2° to the right and, when the saccade to the target started, the target was extinguished; if the saccade amplitude was within a specified range that depended on the recent saccadic history, we delivered reinforcement after that saccade. In *experiments 1, 2, 4*, and *5*, the reinforcement was an auditory tone, whereas in *experiment 3* the reinforcement was the reappearance of the target on the fovea (to avoid having a postsaccadic visual error signal present that depended on the saccade landing position). The reinforcement criteria were computed based on a targeted percentile reinforcement procedure (Galbicka 1994; Galbicka et al. 1993) similar to the one we have used to shape the distributions of saccadic latencies (Madelain et al. 2007). This procedure uses reinforcement and extinction in combination with a systematically changing set of response criteria to reinforce eye movements differentially (i.e., to reinforce saccades meeting an amplitude criterion). It involves two simultaneously operating schedules controlling the amplitudes above and below the goal amplitude (i.e., the saccade amplitude to be learned) that are reinforced. Specifically, after ranking the amplitude of the 50 most recent saccades, we set the lower bound of the saccades to be reinforced to the 10th closest amplitude below the goal amplitude and the upper bound to the 10th closest amplitude above the goal amplitude. By reinforcing only that fraction of saccades ($\pm 20\%$ in our experiments) closest to the goal amplitude, we expected the subsequent saccade amplitude distribution to peak near the goal amplitude. Figure 1A shows the method for a single trial: in this case, the goal amplitude was set to 9.4° (circle). We defined the lower criterion of 8.6° (downward triangle) by the lowest amplitude of the 10 saccades (20% of the previous 50 saccades) smaller than the goal amplitude that were closest to it, and we defined the upper criterion of 10.3° (upward triangle) as the greatest amplitude of the 10 saccades larger than the goal amplitude that were closest to it. On the subsequent trial, only a saccade with amplitude between 8.6 and 10.3° would be reinforced. (In cases in which there had not been 10 trials of the previous 50 saccades below the goal amplitude, all saccades smaller than the goal amplitude would be reinforced. Similarly, if there had not been 10 trials above the goal amplitude, all saccades larger than the goal amplitude would be reinforced. Note that, in either of these cases, fewer than 40% of the previous 50 saccades would be included.) For the first session, the limits were computed based on the cumulative distribution of saccade amplitude

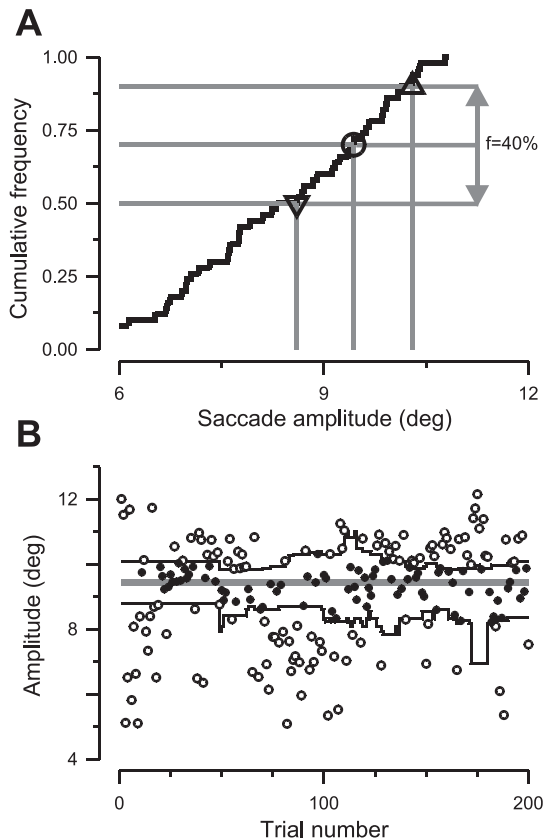


Fig. 1. Reinforcement procedure. *A*: computation of reinforcement criteria for a single trial. Solid black line, cumulative frequency distribution of saccade amplitudes for the previous 50 trials; gray solid lines, corresponding amplitudes and frequency from the cumulative distribution. Circle, goal amplitude; upward and downward triangles, upper and lower limits, respectively, of the reinforcement criterion. *B*: changes in reinforcement criteria over the course of a block (200 trials) for subject S2 (experiment 3). Upper and lower thin black lines, upper and lower limits of the reinforcement criteria, respectively; solid gray line, goal amplitude; unfilled and filled circles, unreinforced and reinforced saccades, respectively.

obtained in the immediately preceding unreinforced baseline session; at the start of subsequent sessions, it was based on the preceding session. This was done because we observed that the adaptation carried over from one session to the next. By basing the reinforcement criteria on the previous session, we ensured that a substantial fraction of saccades would be reinforced. For example, in the experiment shown in Fig. 1*B*, during the reinforced adaptation phase the target step was 14.2° , and we reinforced responses around a value of 9.4° . Notice that both the criteria and hence the amplitude of reinforced saccades changed continuously during the reinforced adaptation, depending both on the goal amplitude and on the amplitudes of the 50 previous saccades.

This procedure has several advantages. First, on any trial, there is a possibility of being reinforced. Second, the overall probability of reinforcement is defined by the experimenter and remains nearly constant throughout the session. Third, once the target behavior is shaped, it reinforces a reduction of the variability of the responses, helping to firmly establish the behavioral pattern. However, because the probability of reinforcement is constant, subjects could adopt the strategy of maintaining the distribution of responses constant. To encourage subjects to alter their saccade amplitudes, an auditory feedback (a synthetic voice saying "Bravo!") indicated that the subject had won one point when three of five consecutive trials were reinforced, including the fifth trial. Subjects were instructed to win as many points as they could as a game. This encourages the saccade

amplitudes to change in the direction of the goal because maximizing the local rate of reinforcement would result in constraining the criteria around the goal amplitude. To increase the reinforcing value of the points, subjects were informed of the number of points earned at the end of each session and knew how many points were earned by the other subjects. Subjects were instructed to try to earn more points than the other subjects.

Because the conventional definition of a reinforcer as an event that has the ability to increase the probability of a particular behavior (Skinner 1981) is circular, the reinforcing value of an event must be experimentally tested. In the present case, one might argue that our "points" are not reinforcers. However, we have successfully used such points to reinforce saccadic latency distributions in previous experiments (Madelain et al. 2007), and the results from the present experiments indicate that giving points can control the subject's behavior and therefore must have some appetitive value. This effect on the subsequent behaviors is also central to distinguish a reinforcer from feedback, defined as a signal that depends on the movement regardless of its ability to induce behavioral modifications. Finally, we should point out that the reinforcer used in our experiment doesn't convey any information about the direction toward which the behavior should be changed, unlike a visual position error feedback.

Experiments

Experiment 1: auditory reinforcement and amplitude decrease. Experiment 1 tested whether reduced gain could be induced using an auditory reinforcement in the absence of a retinal error signal. Four subjects (subjects S1, S2, S5, and S6) performed this experiment. We first collected 200 regular trials without reinforcement (unreinforced baseline session). Each trial started with a fixation period of 500–1,000 ms during which a fixation cross was displayed 7.1° to the left of the center of the screen against a gray background (luminance: 40 cd/m^2). When the fixation cross was extinguished, the target stimulus, a large luminance-defined Gaussian patch (SD 0.93° , maximum luminance: 62 cd/m^2), appeared 14.2° to the right of the fixation cross for 1,000 ms, and the first saccade made in this interval was recorded. A second reinforced baseline session consisted of 200 trials, in which the target stimulus was blanked at saccade onset, and, if the saccade amplitude fell within the reinforcement criteria, with a goal amplitude of 14.2° (the same amplitude as the target eccentricity in the regular trials), an auditory tone (a 60-ms 1,000-Hz sine wave) was delivered as a reinforcement. For the first 50 trials of these baseline sessions, the limits were computed based on the saccade amplitudes obtained during the previous unreinforced baseline session. The subsequent reinforced adaptation sessions differed from the reinforced baseline session in that the goal amplitude was set to 9.4° (66% of the initial target eccentricity). After these adaptation sessions, subjects performed one to three sessions of reinforced recovery trials, similar to the baseline trials, with the destination amplitude again set to 14.2° . The number of reinforced adaptation sessions and reinforced recovery sessions varied among subjects and among experiments. In all experiments, subjects usually performed two sessions of the same adaptation experiment per day, with a 10- to 20-min break between sessions, during which the subject was free to move. If subjects participated in more than one adaptation experiment, at least 7 days elapsed between the experiments.

Experiment 2: auditory reinforcement and amplitude increase. In experiment 2, we set the goal amplitude to 18.9° , 33% larger than the initial target eccentricity, to induce a gain increase in the reinforced adaptation sessions. Except for the goal amplitude, the other parameters were identical to experiment 1. Four subjects (subjects S21, S9, S10, and S11) performed this experiment, of whom only two (subjects S9 and S21) performed the reinforced recovery sessions.

Experiment 3: target view as reinforcement and amplitude decrease. Experiment 3 tested whether the view of the target in the absence of a retinal error signal would induce a decrease in gain. The

design was identical to the one used in *experiment 1* except that, if the saccade amplitude fell within the reinforcement criteria (goal amplitude: 14.2° , the same amplitude as the target eccentricity in the regular trials), the target stimulus then reappeared on the fovea, where it remained for the duration of the trial (see below for details on target stabilization). If the first saccade amplitude did not match the reinforcement criteria, the target remained blanked for the remainder of the trial. Four subjects (*subjects S1, S2, S3, and S4*) performed the experiment.

Experiment 4: auditory reinforcement plus target view and amplitude decrease. In *experiment 4*, we asked whether the auditory reinforcement would be effective despite the presence of the image of the target on the fovea. This experiment used postsaccadic auditory reinforcement as in *experiment 2* except that the target always reappeared on the fovea at the target offset. Here, the auditory tone, but not the vision of the target, depended on the saccade amplitude. Four subjects (*subjects S1, S2, S7, and S8*) performed this experiment.

Experiment 5: comparison between conventional and reinforcement adaptation. To compare some features of adaptation induced by reinforcement and by a conventional double-step paradigm, we had subjects do each paradigm a week apart, and we compared the transfer of the adaptation to other amplitudes and positions and compared the rate of recovery. All subjects completed the reinforcement adaptation first and then the conventional double-step paradigm.

In both cases, the experiment started with a baseline of 200 trials in which the target remained stationary and visible after appearing 14.2° to the right of the fixation point at 7.1° to the left of center. Next, the gain of the saccades was reduced by adaptation: in the case of reinforcement adaptation, there was a session of 200 baseline trials with reinforcement followed by 6 reinforced adaptation sessions. The reinforcement in these sessions was an auditory tone combined with postsaccadic placement of the target on the fovea on every trial as in *experiment 4*. In the case of conventional adaptation, the target stepped back by 4.8° during the saccade for 2 sessions of 200 trials. After the adaptation by either paradigm, there was a transfer session, consisting of 100 more trials of adaptation, followed by 150 trials, during which the target was turned off at the saccade onset. In these trials, we varied either the initial position of the fixation target ("position" trials) or the eccentricity of the target ("amplitude" trials). In the position trials, the initial fixation target was displayed 4.7 or 9.4° to the left of the center of the screen (in addition to the usual 7.1°), and the target appeared 14.2° to the right of it. In the amplitude trials, the fixation position was 7.1° , and the target appeared either 11.8 or 16.5° (in addition to the usual 14.2°) to the right of it. These 5 trial types were presented 30 times each in a pseudorandom sequence. Finally, subjects performed a recovery session in which the first 100 trials were like the preceding adaptation trials, followed by 150 recovery trials, in which the target appeared 14.2° to the right of the fixation point and did not step back during the saccade. Four subjects (*subjects S12, S13, S14, and S15*) performed this experiment.

In addition to the conventional adaptation sessions just described, three other subjects (*subjects S3, S16, and S17*) performed conventional decreased gain adaptation sessions, as described above, and three subjects (*subjects S18, S19, and S20*) performed conventional increased gain adaptation sessions, which were identical to the experiment just described except that in the adaptation trials the target reappeared 18.9° away from the fixation point at the saccade offset, thereby requiring a gain increase.

Participants

Of the 21 subjects, 19 subjects were naïve with respect to the goal of the experiments and had no previous experience in oculomotor experiments. In addition, two of the authors (*subjects S1 and S2*) were the subjects of *experiments 1, 3, and 4*. As indicated in the RESULTS, some subjects participated in more than one experiment. To familiarize the naïve subjects with the eye movement recording apparatus and

calibration procedures, they first performed a session of 200 trials in which they made saccades to the appearance of a target 14.2° to the right from the fixation point. We do not report results from these familiarization trials. All participants had normal vision or vision corrected to normal. All experimental procedures were reviewed and approved by the Institutional Review Board, and each subject gave informed consent.

Apparatus

Stimuli were generated on a Power Mac G4 using the Psychophysics Toolbox extensions for Matlab (Brainard 1997; Pelli 1997) and displayed on a video monitor (Iiyama HM204DT, 100 Hz) at a viewing distance of 60 cm. To minimize measurement errors, the subject's head movements were restrained using a dental impression and a forehead rest, so that the eyes in primary gaze position were directed toward the center of the screen. The presentation of stimuli and acquisition, display, and storage of responses were controlled by a PXI computer using the LabView Real-Time software package (National Instruments). Trigger signals from the visual display computer to the LabView computer allowed us to synchronize data collection to stimulus presentation with 1-ms resolution. Eye movements were measured continuously with an infrared video-based eye tracking system (ISCAN, RK-726) at 240 Hz. Before each experimental session, we calibrated the eye tracker by having subjects repeatedly fixate a set of seven horizontal locations to generate a linear function for converting raw eye tracker values to horizontal eye position.

Online Saccade Detection

For online saccade detection, a real-time algorithm used a velocity threshold to identify the start and end of the first saccade that occurred within 1000 ms after the target appeared (saccades occurring before the target appeared or 1,000 ms after it were not detected). In *experiments 3–5*, the target was blanked at the saccade onset and reappeared at the fovea at the end of the saccade and stayed at that position. To do this, we determined the median eye position over 12 ms after the saccade and then had the target reappear at this location as soon as the position computation was completed. By comparing the online measurement used to place the target on the fovea with offline measures of postsaccadic eye position, we found that the target was placed, on average, 0.17° (SD 0.31°) ahead of the eye position. Upon the saccade onset, the target was blanked for an average of 65 ms.

Data Analysis

For offline analysis of the eye movement recordings, an interactive analysis program was used to filter, display, and analyze the data. Horizontal eye velocities and acceleration were obtained by differentiating the eye position signal. Saccades were detected by applying a set of fixed velocity and acceleration criteria (30% and $7,000\%$, respectively) using a purpose-written LabView software that presented the start and end of each saccade to the investigator for confirmation or adjustment. Saccade amplitudes were computed by subtracting the median eye positions across 25-ms windows before and after the saccade. We measured only the first saccade within 1,000 ms after the target onset. We omitted trials in which the saccade was opposite in direction to the target or had a latency shorter than 100 ms.

To show the trends of the reinforced adaptation and reinforced recovery in individual experiments, we smoothed the records of saccade amplitude versus trial number by means of a Lowess iterative nonlinear regression (width: 100 trials) and applied it separately for each session, so that we did not smooth across session boundaries. To describe the changes in amplitude, we computed the mean amplitude of the last 50 trials in each session. To quantify the extent of change in gain, we compared the 50 trials at the end of the reinforced baseline

session (in which the goal amplitude had been set to 14.2°) with the 50 trials at the end of the last reinforced adaptation session (goal amplitude either 9.4 or 18.9°), yielding the percentage of amplitude change as $100 \times [(\text{mean reinforced adaptation amplitude} - \text{mean baseline amplitude}) / \text{mean baseline amplitude}]$. Because the target always stepped by the same amount, this percentage is equivalent to the proportion of gain change. This measure is more conservative than the raw saccadic gain at the end of the reinforced adaptation, because it takes into account the changes in saccade amplitude during the baseline sessions. We also computed the percentage of amplitude change in the last 50 reinforced recovery trials with respect to the last reinforced adaptation session as $100 \times [(\text{mean amplitude in recovery} - \text{mean amplitude in adaptation}) / \text{mean amplitude in adaptation}]$.

RESULTS

Overall, we found that reinforcement, by either an auditory signal or a visual signal, could change the gain of saccades in a progressive manner.

Experiment 1: Auditory Reinforcement

Figure 2A shows the saccadic gains for each trial in one subject (*subject S6*). In the unreinforced baseline session, the mean saccadic gain was 0.93, well within the range of gain reported in the literature. The gain reached a mean value of 0.94 at the end of the reinforced baseline session in which the reinforcement criteria were set to maintain a gain close to unity. During the reinforced adaptation sessions, the amplitude diminished steadily, reaching a mean gain of 0.70 in the last 50 adaptation trials with hardly any loss of adaptation from one day to the next (heavy vertical gray line at *trial 800*). In this subject, the gain reached the goal amplitude by the fourth reinforced adaptation session. During the reinforced recovery sessions, in which the goal amplitude was set back to 14.2°, the recovery was only transient.

All subjects showed a consistent decrease in saccade amplitude during the reinforced adaptation sessions (Fig. 2B). Two of the subjects had reached the goal amplitude by the end of the second reinforced adaptation session, whereas *subjects S5* and *S6* had the lowest amplitude in the sixth and fourth reinforced adaptation session, respectively. The proportion of amplitude change from baseline to adaptation trials was -23.1% on average. The amplitude during reinforced recovery generally increased except for *subject S6*, in which the saccade amplitude decreased during the last recovery session.

To evaluate the reliability of the change in saccadic gain caused by the reinforcement procedure, we compared the gains obtained in the last 50 trials of unreinforced baseline, baseline, last reinforced adaptation, and last reinforced recovery sessions using individual one-way ANOVA and post hoc pairwise two-tailed *t*-tests using the Dunn-Sidak correction (Fig. 2C). The session type had a significant effect in all subjects [ANOVA (3,196), $P < 0.05$], and post hoc pairwise tests (Fig. 2C, brackets on top of the histograms) further revealed a decrease in gain during the adaptation session and an increase during the reinforced recovery session, although one subject (*subject S6*) did not show a significant increase from the adaptation session to the reinforced recovery session. Finally, by comparing the saccades made to simple target steps without reinforcement (unreinforced baseline) to those made to simple target steps with reinforcement (baseline condition), we found that the gain decreased significantly in two subjects (*subjects S2* and *S5*, corrected *t*-test, $P < 0.05$).

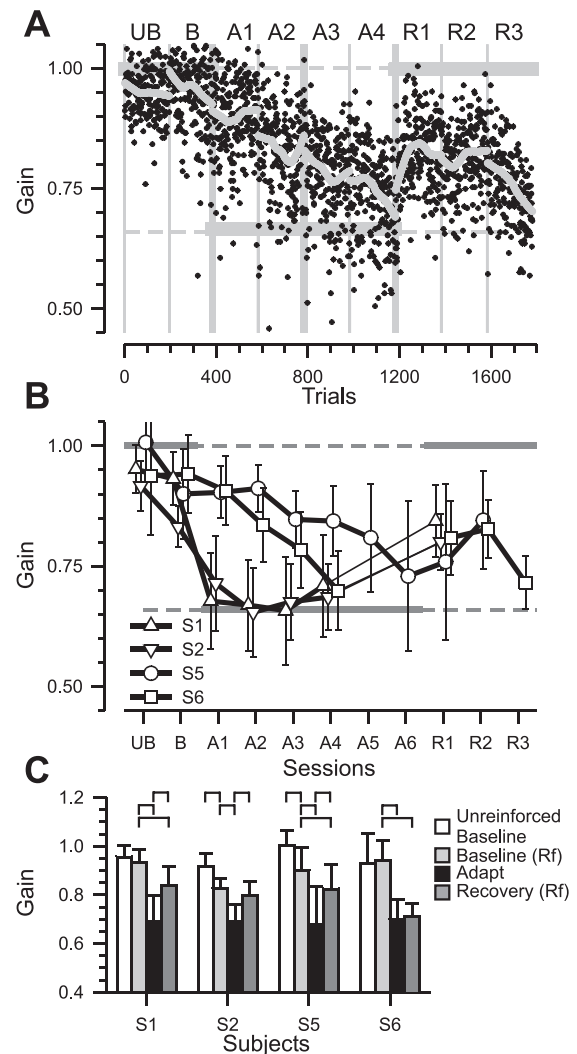


Fig. 2. Decreased gain auditory reinforcement (experiment 1). A: changes in saccadic gain for a single subject (*subject S6*). Solid circles, saccadic gains; solid gray curves, saccadic gain fit by Lowess smoother. Solid horizontal gray lines, goal amplitude; vertical gray lines separate sessions [letters on top indicate the session type: unreinforced baseline (UB), reinforced baseline (B), reinforced adaptation (A), and reinforced recovery (R)]; thick gray lines separate days. B: mean (\pm SD) saccadic gain across the last 50 trials of each session for four subjects. Thin lines connect the last adaptation session with the first reinforced recovery session. Solid horizontal lines, goal amplitude. C: bar graphs plot mean (\pm SD) saccadic gain for the last 50 trials of each session for each subject. Horizontal brackets indicate comparisons that were significant by post hoc pairwise *t*-tests ($P < 0.05$).

Experiment 2: Increased Gain Adaptation with Auditory Reinforcement

In this experiment, reinforcement contingencies were set to encourage increased saccade amplitude. Figure 3A shows the time course of saccadic gain for one subject (*subject S9*). During the preadaptation and baseline sessions, gains were 0.98 and 0.92, respectively. A marked increase in saccade amplitude then occurred, starting during the fourth reinforced adaptation session, such that the gain at the end of adaptation reached 1.07. For *subject S9*, the proportion of amplitude change was 19.7% and saccade amplitude decreased during the reinforced recovery sessions (-12.9%). All subjects but *subject S10* increased their gain during the adaptation sessions.

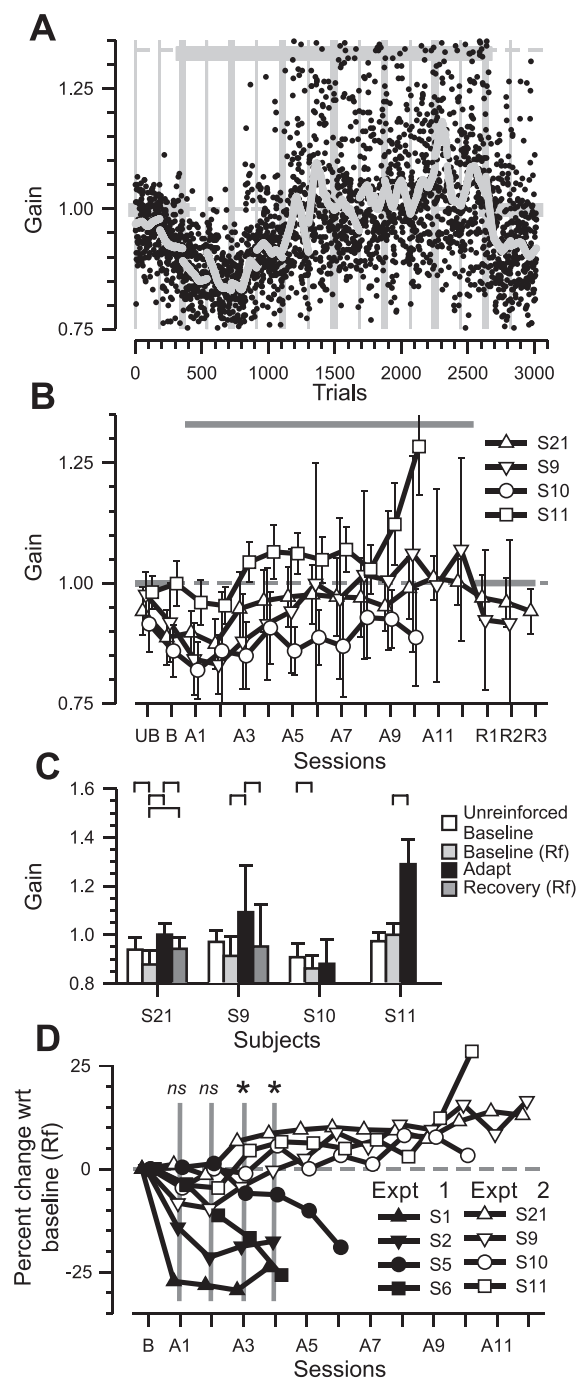


Fig. 3. Increased gain auditory reinforcement (experiment 2). A: changes in saccadic gain for a single subject (subject S9). B: mean (\pm SD) saccadic gain across the last 50 trials of each session for four subjects. C: mean (\pm SD) saccadic gain for each subject in each condition. Symbols and lines are as in Fig. 2. D: percent change during reinforced adaptation in experiment 1 (filled symbols) and experiment 2 (open symbols). Dotted horizontal line, no change. Vertical lines with asterisks indicate comparisons that were significant by rank sum tests ($P = 0.029$). ns, Not significant.

SDs were small in most subjects (0.06, 0.08, and 0.06 in subjects S21, S10, and S11, respectively; Fig. 3B), with the exception of subject S9 (0.15 in gain units on average during the reinforced adaptation sessions), more than twice as much as in the other subjects, suggesting the possibility that this subject was probing the limits of the reinforced amplitude range. We

found a significant effect of the session type [ANOVA (3,196), $P < 0.05$] in three of four subjects. Post hoc pairwise tests revealed a significant (corrected t -test, $P < 0.05$) increase in gain in these subjects (subjects S21, S9, and S11) when we compared the baseline and reinforced adaptation conditions (Fig. 3C). In subjects S9 and S21, who experienced a reinforced recovery condition, the gain significantly decreased in the reinforced recovery condition compared with the reinforced adaptation condition; the gain after the recovery was not different from the baseline gain.

Figure 3D shows the percentage of amplitude change in the reinforced adaptation session with respect to the baseline amplitude for experiments 1 and 2. It should be noted that the only differences between these experiments were the goal amplitudes during the reinforced adaptation sessions (9.4° and 18.9° , respectively), which were used to compute the reinforcement criteria. The results indicate that the amplitude followed the reinforcement criteria in that the amplitude decreased in experiment 1 (downward adaptation, average percent change: -23.1%) and increased or remained unchanged in experiment 2 (forward adaptation, average percent change: $+16.2\%$). Percentages of change from each experiment were compared in the first four sessions using a rank-sum test: although no significant differences were observed in the first and second adaptation session, the gain changes in experiments 1 and 2 were significantly different during the third and fourth adaptation sessions ($P = 0.029$). This indicates that the changes in amplitude were in fact induced by reinforcement and not by some passive change in gain, a point we will further discuss in the DISCUSSION.

Experiment 3: Target View as Reinforcement

When we blanked the target at the saccade onset and presented a postsaccadic target view after saccades that met our criteria, changes in gain were also obtained. The time course of change in gain for one subject (subject S4) is shown in Fig. 4A. Notice that in the last 50 trials of the baseline session, the gain decreased continuously (see Lowess curve) and reached a mean value of 0.82. During the reinforced adaptation sessions, the amplitude diminished steadily, reaching a mean gain of 0.67 in the last 50 adaptation trials. Notice that the gain recovered partially between sessions and quickly reached a value at or below that of the previous session, even when a day elapsed between the sessions (see Lowess curve). In this subject, the gain reached the goal amplitude by the second reinforced adaptation session. During the reinforced recovery sessions, in which the goal amplitude was set back to 14.2° , the amplitude progressively increased to reach a mean value of 0.80 in the last 50 trials.

Across the four subjects, the overall pattern appeared similar to the one from experiment 1 (Fig. 4B). The proportion of amplitude change from baseline to adaptation trials was -21.7% on average. The amplitude during reinforced recovery significantly increased in all subjects. The session type had a significant effect in all subjects [ANOVA (3,196), $P < 0.05$], and post hoc pairwise tests (Fig. 4C) further revealed a pattern similar to the one observed in experiment 1: a decrease in gain during the adaptation session and an increase during the reinforced recovery session.

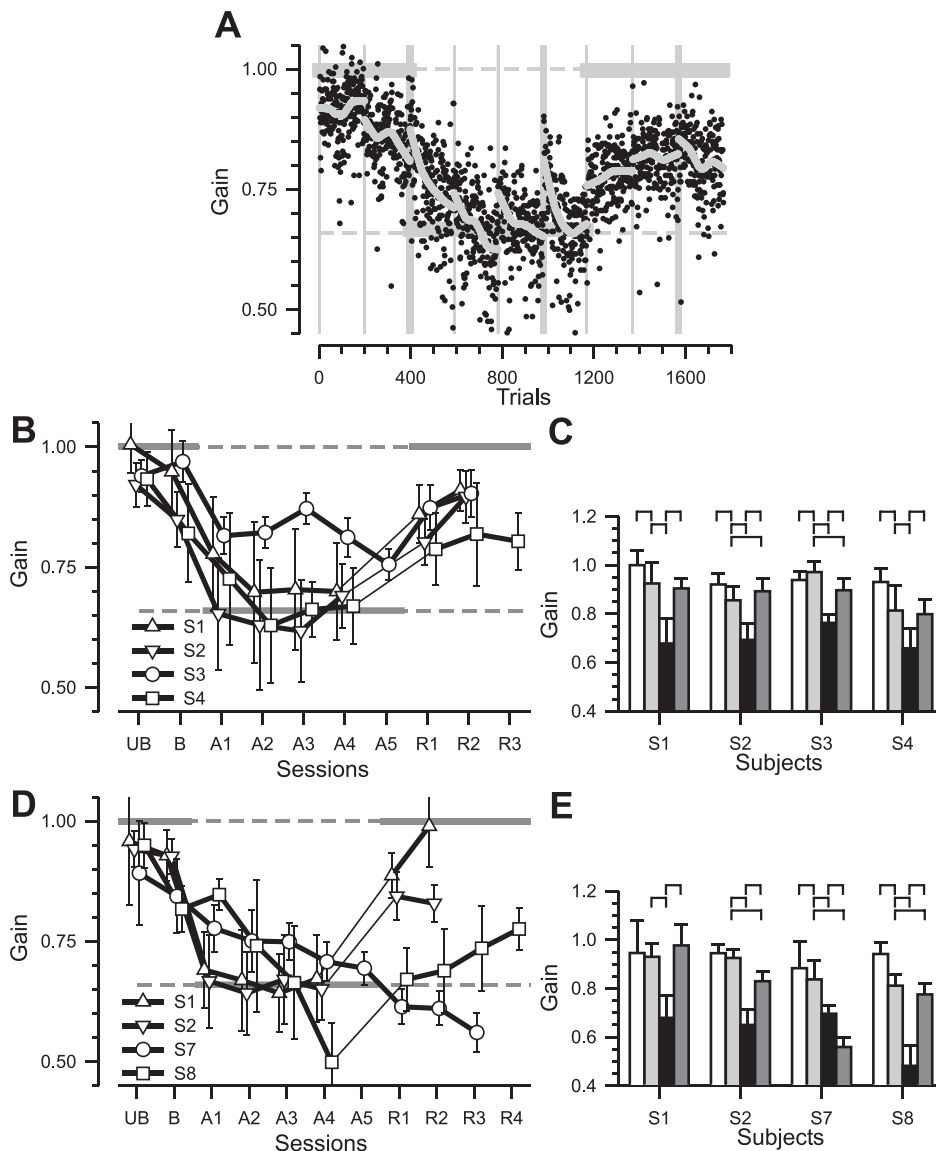


Fig. 4. Target view reinforcement and auditory reinforcement with target view after every saccade (experiments 3 and 4). *A*: changes in saccadic gain for a single subject (subject S4) during experiment 3 (target view reinforcement). *B*: mean (\pm SD) saccadic gain across the last 50 trials of each session of experiment 3 for four subjects. *C*: mean (\pm SD) saccadic gain for each subject (experiment 3). *D*: mean (\pm SD) saccadic gain across the last 50 trials of each session of experiment 4 (auditory reinforcement with target view after every saccade) for four subjects. *E*: mean (\pm SD) saccadic gain for each subject (experiment 4). Symbols and lines are as in Fig. 2.

Experiment 4: Auditory Reinforcement Plus Target View

When auditory reinforcement was presented after saccades that met our criteria, but the target view was presented after all saccades, subjects showed a gain decrease as strong as when the auditory reinforcement was presented alone. We found a marked reduction in amplitude in all subjects with high percentages of amplitude change (-28.5% on average; Fig. 4*D*). Saccade amplitude increased during the reinforced recovery sessions in three subjects but continued to decrease in subject S7.

We found a significant effect of the session type in all subjects [ANOVA (3,196), $P < 0.05$], and post hoc pairwise tests (Fig. 4*E*) revealed a pattern similar to the one observed in experiments 1 and 2. One subject (subject S7) showed a significant decrease from the reinforced adaptation to the reinforced recovery session.

Experiment 5: Comparison Between Conventional and Reinforced Adaptation

To compare some key characteristics of reinforced and conventional adaptation, we had four subjects do reinforced

adaptation (auditory stimuli as reinforcement and target view after every saccade) followed a week later by conventional adaptation. All the subjects sequentially tested on reinforced and conventional adaptation showed a consistent decrease in saccade amplitude during both the reinforced and conventional adaptation sessions (Fig. 5, *A* and *B*). Across subjects, the saccadic gain reached an average of 0.67 for the reinforced adaptation and 0.68 for the conventional adaptation (the goal was 0.66), which represented, on average, -22.8% and -27.4% of amplitude change, respectively.

To evaluate the changes in saccadic gain in both paradigms, we compared the gains as we did in the other experiments: we used individual one-way ANOVA [ANOVA (6,343), $P < 0.05$ in all four subjects] and post hoc pairwise two-tailed *t*-tests using the Dunn-Sidak correction to compare the last 50 trials of the unreinforced baseline, reinforced baseline, last reinforced adaptation, first recovery, conventional baseline, conventional adaptation, and second recovery sessions (Fig. 5*B*). Post hoc pairwise comparisons revealed reinforcement-induced changes similar to what we observed in the other reinforced experi-

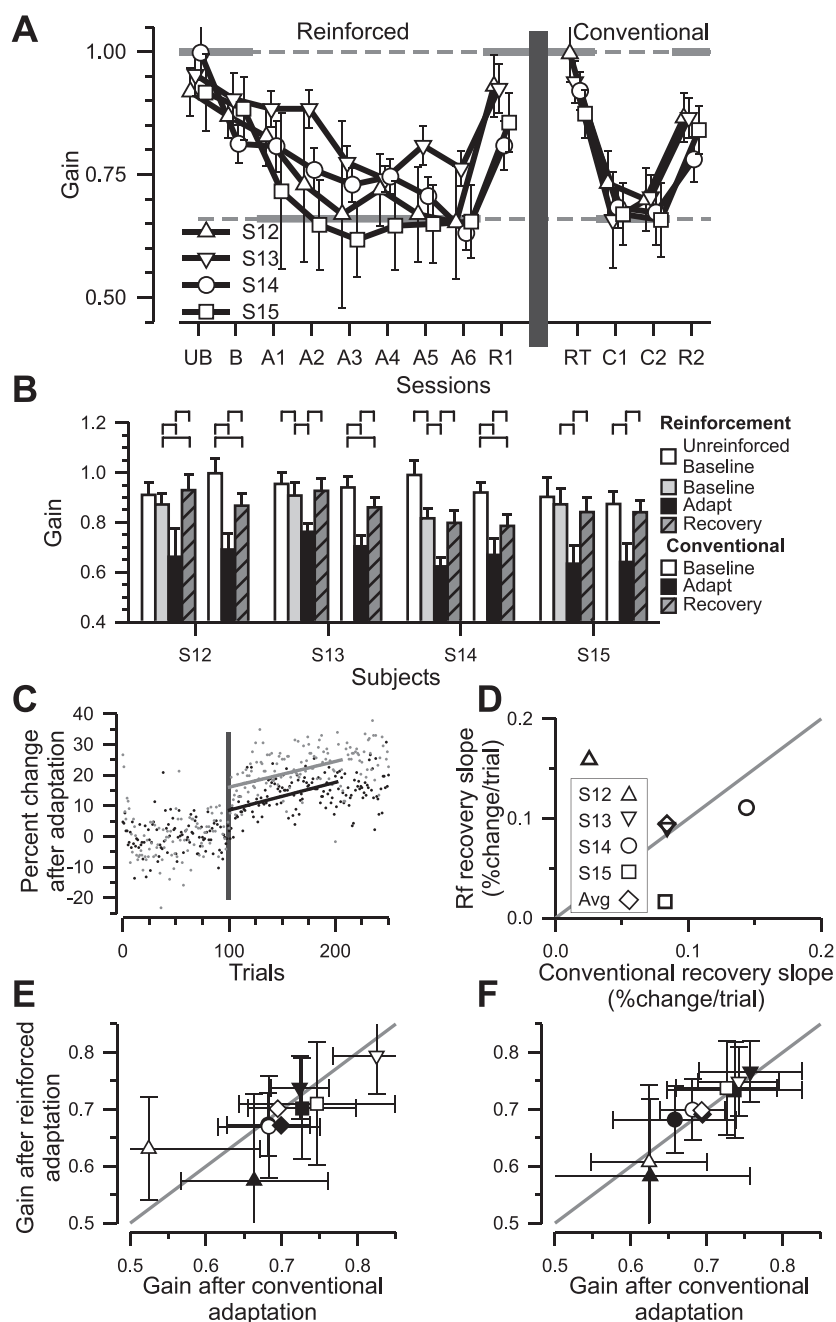


Fig. 5. Comparison of reinforced and conventional adaptation (*experiment 5*). **A**: mean (\pm SD) saccadic gain across the last 50 trials of each session of *experiment 5* for four subjects (transfer sessions not shown). The thick vertical gray line indicates the week separating the reinforced adaptation experiment and the conventional adaptation experiment. Symbols and lines are as in Fig. 2B. **B**: histograms of the mean (\pm SD) saccadic gain for the last 50 trials of each session (for each subject, first 4 columns are reinforced adaptation and last 3 columns are conventional adaptation). The hatched bars show unreinforced recovery sessions, in which the target was always visible and did not move upon saccades. **C**: percent change during conventional recovery for a single subject (*subject S13*). Solid circles and gray line, saccadic gain in the reinforced adaptation condition; shaded circles and gray line, saccadic gain in the conventional adaptation condition. The vertical gray line indicates the beginning of recovery trials. **D**: recovery slopes (reinforced vs. conventional adaptation). **E**: gain in the position transfer trials (reinforced adaptation vs. conventional adaptation). Solid symbols, 9.4° initial fixation; open symbols, 4.7° initial fixation. Symbols are as in **D**. Error bars indicate SDs. **F**: gain in the amplitude transfer trials (reinforced adaptation vs. conventional adaptation). Solid symbols, 16.5° target eccentricity; open symbols, 11.8° target eccentricity. Symbols are as in **D**. Error bars indicate SDs.

ments as well as the expected changes in the conventional paradigm. When similar sessions from each paradigm were compared, the gain was significantly lower in the reinforced baseline condition than in the conventional baseline in three subjects (not in *subject S15*); compared with the conventional adaptation session, the reinforced adaptation had a lower gain in one subject (*subject S14*) and a higher gain in another subject (*subject S13*), and there was no difference in two subjects (*subjects S12* and *S15*). Finally, recovery from the reinforced adaptation was greater than from the conventional adaptation in two subjects (*subjects S12* and *S13*) but not different in the other two subjects. (In both recovery sessions, the target remained visible and did not move during the saccade.)

To compare the rate of recovery after reinforced adaptation and after conventional adaptation in the same subjects, we computed the trial-by-trial gain as a percent increase from the mean gain obtained in the first 100 trials of the adaptation session $\{100 \times [(\text{recovery trial gain} - \text{mean adapted gain}) / \text{mean adapted gain}]\}$, where 0% indicates no recovery and 100% indicates a doubling of the adapted gain. The first 100 recovery trials were then fitted using a linear regression. (Figure 5C shows an example for one subject.) The average rates of recovery were higher after conventional than reinforced adaptation for two subjects (*subjects S14* and *S15*) and lower for the other two subjects (*subjects S12* and *S13*), as shown in Fig. 5D. All subjects showed significant increases in gain from the adaptation condition to the recovery condition; in three of

the four subjects, the gain during the recovery session was not significantly different from that in the baseline session (Fig. 5B).

The time course of adaptation was more variable in the reinforced adaptation than in the conventional adaptation. In the conventional adaptation sessions, the adaptation was essentially complete within the first session in all subjects (i.e., the gain was close to that demanded by the postsaccadic position of the target). In the reinforced adaptation sessions, the gain became lower than 0.7 (target gain was 0.66) by the end of the second session for *subject S15*, by the end of the third session for *subject S12*, and by the end of the sixth session for *subject S14*.

When we tested the adapted subjects on target steps of different amplitude or different starting positions, we found that both the reinforced adaptation and conventional adaptation transferred to a similar degree (Fig. 5, *E* and *F*): position transfer trials, gains were 0.67 and 0.70, respectively, for the 9.4° starting point and 0.70 and 0.69 for the 4.7° starting point; and amplitude transfer trials, 0.69 and 0.69 for the 16.5° target eccentricity and 0.70 and 0.69 for the 11.8° eccentricity. The only significant differences were observed in *subject S12*, in whom the gain at the 9.4° starting point was significantly smaller after reinforced than conventional adaptation (0.57 vs. 0.66, $P = 0.01$) and significantly larger after reinforced than conventional adaptation at 4.7° (0.62 vs. 0.52, $P = 0.001$). After both types of adaptation, the gains on all transfer trials were within 5% of the gains after the adaptation. This almost perfect transfer of adaptation in both paradigms implies that the adaptation was not an artifact of a strategic targeting of saccades to specific locations.

In the subjects tested only with conventional adaptation (Fig. 6, bottom two groups of bars), those tested with the decreased

gain adaptation showed a decrease in gain [ANOVA (3,196), $P < 0.05$ in all subjects], with post hoc pairwise tests revealing a significant decrease in gain from the baseline to adaptation sessions in all three subjects (from an average of 0.94 to 0.72, an amplitude change of -22.6% , corrected t -test, $P < 0.05$); the three subjects tested with the conventional increased-gain adaptation [ANOVA (3,196), $P < 0.05$ in all subjects] showed a 13% increase in gain to 1.02 from the 0.90 recorded during the baseline sessions (corrected t -test, $P < 0.05$), which was less than the decrease of -22.6% after the decreased gain adaptation, although the intrasaccadic step of the target was 33% for 400 trials in both cases. We also observed a recovery in both the decreased gain experiments (average amplitude change: 23.7% , corrected t -test, $P < 0.05$) as well as increased gain experiments (average amplitude change: -6.3% , corrected t -test, $P < 0.05$). However, the gains were still significantly lower in the recovery condition than in the baseline condition (corrected t -test, $P < 0.05$).

General Observations on Results From Reinforced Adaptation Experiments

All three types of decreased gain reinforcement (target view, auditory, and auditory with target view after all saccades) used in *experiments 1* and *3–5* resulted in a substantial reduction in gain (Fig. 6). The amount of adaptation was similar [ANOVA (3,3), $P = 0.42$] in the auditory condition (23% in *experiment 1*), target view condition (22% in *experiment 3*), and auditory reinforcement with every saccade followed by target view condition (28% in *experiment 4* and 23% in *experiment 5*). Overall, the gain was reduced by an amount similar to that shown in the conventional (intrasaccadic back step) adaptation sessions (24% vs. 25%).

When the reinforcement was switched to encourage larger saccades during the reinforced recovery conditions (*experiments 1, 3, and 4*), the gain increased almost to the baseline levels (average amplitude change: 21.7%) in all but two subjects (*subjects S6 and S7*). There was no substantial difference in the degree of recovery across reinforcement conditions. During the reinforced gain increase experiment (*experiment 2*), the gain increased by 16.2% compared with the baseline session. In comparison, the conventional increased gain adaptation experiment showed a 13% increase (Fig. 6). As was the case with the decreased gain experiments, when saccades back to the original target location were reinforced during the reinforced recovery session, the gain declined to nearly the baseline level.

DISCUSSION

Our results show that, in the absence of retinal error, an arbitrary reinforcement (either an auditory tone or viewing the target on the fovea) can guide changes in saccade amplitude. We will discuss first how reinforcement may relate to saccade adaptation, then the implications of our particular protocols, and, finally, the similarities and differences between saccade adaptation produced by reinforcement and by conventional intrasaccadic target steps.

As mentioned above, saccadic adaptation is classically seen as a servomechanism in which saccadic gain for targets of a particular eccentricity and direction is controlled over many trials by a retinal error signal—the relative position of the target

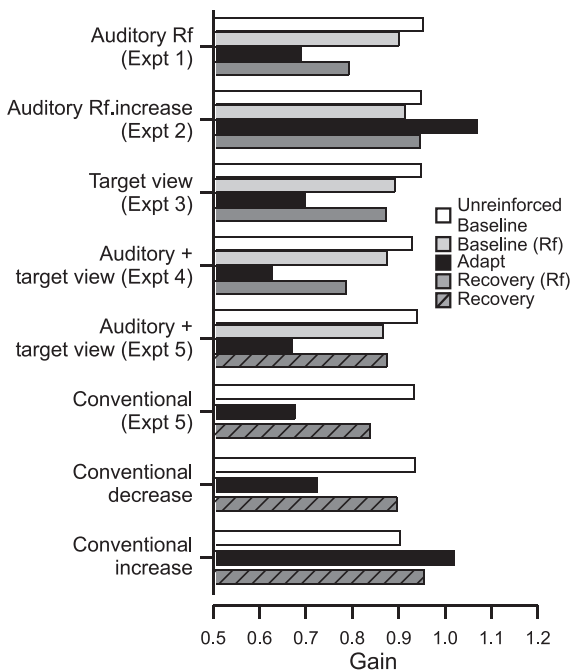


Fig. 6. Summary of experiments. Bar graphs plot the mean saccadic gain for the last 50 trials of unreinforced baseline, reinforced baseline, reinforced adaptation, reinforced recovery, and conventional (unreinforced) recovery averaged across subjects. The lower two sets of bars refer to the six subjects who did only the conventional adaptation.

and fovea after each saccade (Noto and Robinson 2001; Wallman and Fuchs 1998). Were this the only error signal, one would expect that placing the target on the fovea after each saccade would constitute “error clamping” and would prevent adaptation. This, however, is not the case (Havermann and Lappe 2010; Henson 1978; Robinson et al. 2003). A second possible error signal is that the oculomotor system responds to the target being in an unexpected place after a saccade (failure of prediction). Thus, when subjects were told to make saccades two-thirds of the way to the target, backstepping the target during saccades resulted in decreased saccadic gain even though the backstepped target was beyond the landing position of the saccade (Bahcall and Kowler 2000). Similarly, Wong and Shelhamer (2011) and Havermann and Lappe (2010) showed that saccadic gain could be decreased even by extremely small intrasaccadic step backs, which left the postsaccadic target beyond the fovea, a situation that should have produced increased gain if retinal error were the only signal driving adaptation. Here, we introduce a third mode of saccade adaptation: saccadic gain may be controlled by a more general learning process, reinforcement. Because some saccades land closer to the target than others, the oculomotor system could select the movements that had better value while avoiding the ones that had lower value, thereby progressively adjusting its amplitudes during learning to compensate for the target’s back step. In other words, adjusting the saccadic gain could be explained by a selection process (Donahoe et al. 1993) in which those responses that produce a valuable outcome—maximizing visual perception—are selected over those that do not.

A fundamental distinction between the feedback afforded by reinforcement and that afforded by either retinal error or prediction error is that on individual trials reinforcement does not have a sign, whereas the other two error signals indicate on which side of the target the eye has landed after each saccade.

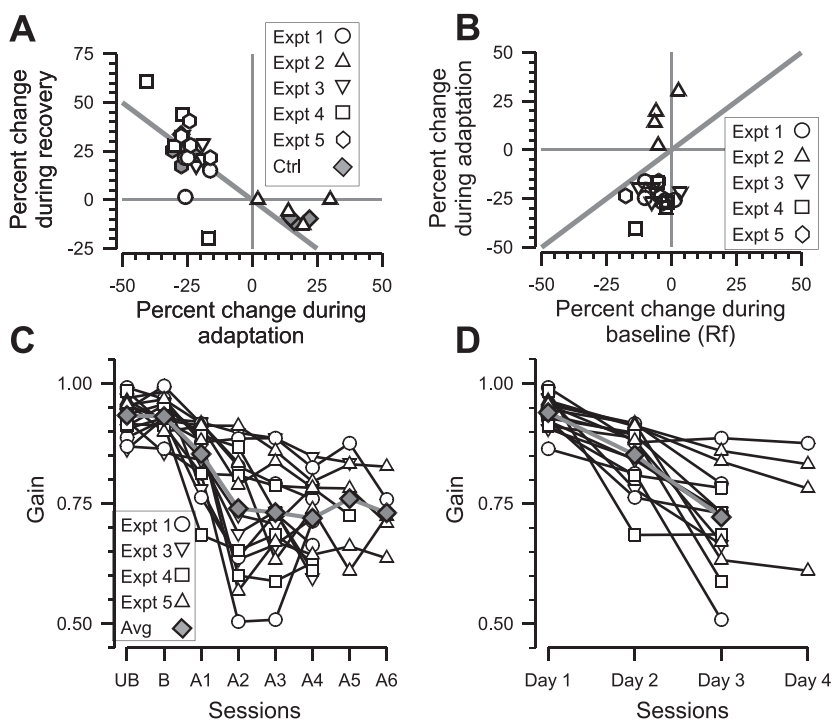
Thus, in our study, *experiments 1* and *2* are “pure” reinforcement experiments in that no directional signal is present.

Our results show that reinforcement is a plausible mechanism for guiding saccade adaptation, but we do not attempt to weigh its importance against that of retinal error or failure of prediction of postsaccadic target location. Rather, we show that the absence of postsaccadic visual position error, and without any postsaccadic visual stimulus, the amplitude of saccades can be manipulated. In particular, in *experiments 1* and *2*, subjects did not have a postsaccadic retinal image of the target but adaptation occurred regardless, directed toward lower or higher gain by an auditory tone. In *experiment 3*, a view of the target was presented on the fovea only after saccades of the desired size. Although the decrease in gain might be explained by the fact that the target had stepped back, resulting in a prediction error, this could not explain the recovery when the same stimulus was presented in response to saccades of increasing amplitude.

Are the Changes in Amplitude Attributable to Reinforcement Contingencies?

There are strong reasons for attributing the changes in saccade amplitude to the schedules of reinforcement that we imposed. First, in the experiments in which decreased gain was reinforced, the gain decreased, whereas in those in which increased gain was reinforced, the gain increased or was unchanged. Second, during the recovery phase of each experiment, when the direction of gain change reinforced was reversed, the changes in gain also reversed in nearly all cases (Fig. 7A). As shown in Fig. 7A, gain decreases during backward adaptation followed by gain increases during recovery would plot in the upper left quadrant of the graph, whereas gain increases during forward adaptation experiments followed by gain decreases during recovery would plot in the lower right quadrant. It is apparent that for all but two of the subjects, the

Fig. 7. *A*: change in gain during recovery as a function of the change in gain during adaptation. Open symbols, reinforced adaptation; shaded symbols, conventional adaptation. *B*: change in gain during reinforced adaptation as a function of the change in gain during reinforced baseline for *experiments 1–5*. *C*: retention of adaptation from session to session for each subject in *experiments 1* and *3–5*. Mean saccadic gain across the first 25 trials of each session is shown. *D*: retention of adaptation from day to day. Mean saccadic gain across the first 25 trials of each session is shown. Symbols are as in *C*.



degree of both gain decrease and gain increase during reinforced recovery was comparable to that of the conventional (intrasaccadic backstep) adaptation experiments (diamonds). This pattern of results indicates that changes in gain were well controlled by the reinforcement contingencies: gain decreased when the contingencies favored smaller saccades and increased when larger saccades were reinforced. Because the type of reinforcement did not have an obvious effect on the degree of change in gain (Fig. 6), we conclude that it is reinforcement, per se, rather than the particular reinforcer used, that accounts for the modification of saccadic gain.

One peculiarity of the reinforcement-guided changes in gain that might undermine our interpretation of the gain decrease results is that when the reinforcement contingencies were set to induce a gain close to unity during the baseline session, the gain decreased in 60% of the cases, although the average gain decreased very little (from an average of 0.94 to 0.89). If this rate had continued throughout all our adaptation sessions, it could have produced a substantial gain decrease independent of the reinforcement contingencies. We are confident that this downward baseline drift does not explain the adaptation we found because the change during adaptation was uncorrelated with those during the baseline period (Fig. 7B). Furthermore, during the adaptation trials, the gain changes were either upward or downward, depending on the reinforcement criteria.

The explanation for the downward drift of the gain during the baseline period may be related to the tendency of normal saccades to have gains below one, a tendency that is actively maintained (Becker 1989; Harris 1995; Henson 1978). One can speculate that in the absence of postsaccadic visual error, the gain has a tendency to drift downward. This effect might be amplified in our reinforcement paradigms. In the case of the auditory reinforcement of *experiment 1*, the blanking of the target would have left no retinal error to counteract for the downward drift, whereas in *experiments 3–5*, in which the target is stabilized on the fovea after the saccade, this would simulate the situation in which saccades were larger than their usual set point and might thereby induce a gain decrease, much as in the experiments of Henson (1978) and Robinson et al. (2003), in which placing the target on the fovea caused a gain decrease.

Adaptation Versus Recovery and Backward Versus Forward Adaptation

As in conventional saccade adaptation experiments (Hopp and Fuchs 2004), we found more robust adaptation in the gain decrease situation than in the gain increase situation, as manifested in *experiment 1* compared with *experiment 2*, as well as in the recovery being incomplete in 5 of the 12 subjects and absent in 2 subjects. (Perhaps if we had given subjects as many recovery sessions as adaptation sessions, the gains would have all returned to baseline.) Indeed, there is evidence that the changes in gain consist of several components with different rate constants (Chen-Harris et al. 2008; Kojima et al. 2004), implying that the gain at the moment does not completely describe the state of the saccade machinery.

The origin of these differences in forward and backward adaptation is unknown but could be related to saccades being naturally hypometric (Henson 1978; Straube et al. 1997),

perhaps because of a tonic tendency for the gain to decrease. It has also been suggested that distinct mechanisms are involved: a reduction of motoric gain in the case of backward adaptation and a remapping of the desired goal for saccades in the case of forward adaptation (Hernandez et al. 2008; Semmlow et al. 1989).

Retention of Adaptation

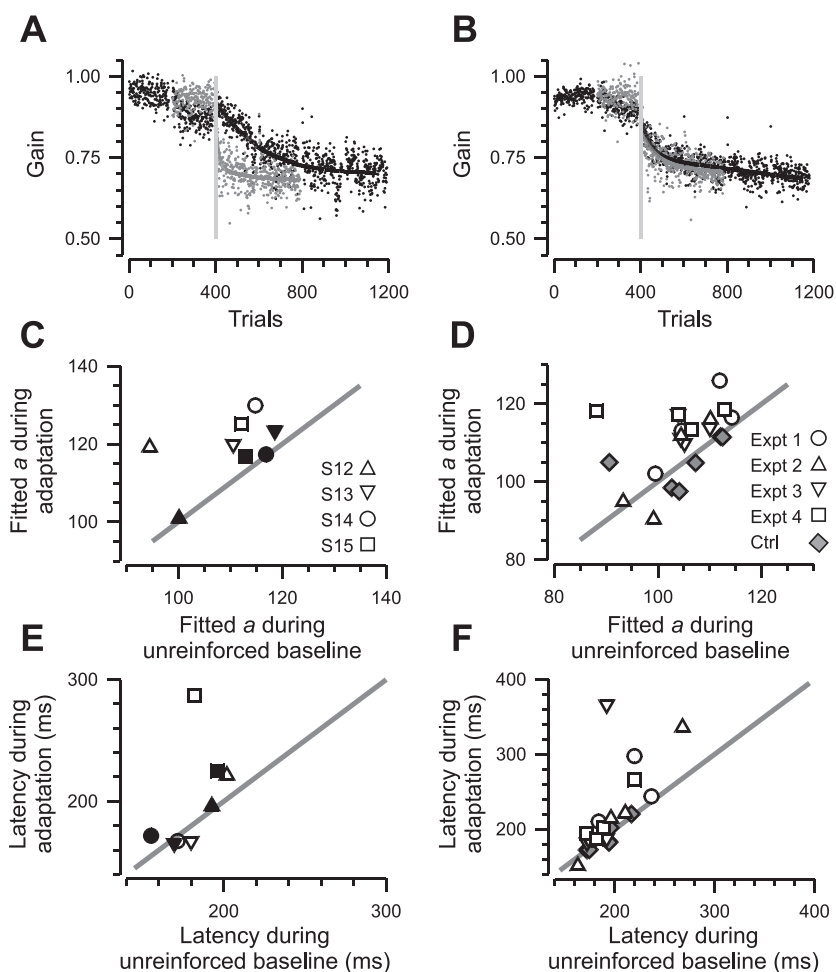
During the decreased gain experiments, the reductions in saccade amplitude were generally maintained across sessions, and even across days, although subjects were, of course, not kept in the dark between the sessions. The mean amplitude of the first 25 trials of each session showed a steady decrease across sessions (Fig. 7C) and even across days (Fig. 7D). Similar retention occurs after conventional adaptation in both monkeys and humans. In monkeys, some maintenance of adaptation has been reported after 24 h under normal illumination (Noto et al. 1999). In humans, the gain decrease induced by a back step protocol using 10–20° target steps was considerable the next day and lasted up to 5 days (Alahyane and Pelisson 2005).

This persistence of change in saccadic gain might have several explanations: although subjects make thousands of unadapted saccades between sessions, few of them would be the “reflexive” saccades to the sudden movement or appearance of a stationary target, and few would be as large as the saccades studied in most saccadic adaptation experiments. The adaptation of reflexive saccades is somewhat independent of adaptation of the “scanning” saccades common to everyday life (Deubel 1999; Fujita et al. 2002; Hopp and Fuchs 2010), and the adaptation of large saccades is somewhat independent of adaptation of smaller saccades (Noto et al. 1999). Furthermore, saccadic adaptation is sensitive to context, in that subjects can maintain more than one gain for identical saccades depending on proprioceptive or visual contexts (Alahyane and Pelisson 2004; Herman et al. 2009; Shelhamer and Clendaniel 2002). Even though we did not explicitly manipulate context in our experiments, the experimental setup necessary to record eye movements, the bite bar, or the small target on a dark screen might constitute a strong enough context for these learned contingencies to yield gains stable across days.

Similarities and Differences Between Reinforcement-Induced Changes in Saccade Amplitude and Conventional Saccadic Adaptation

Although the reinforcement procedures we used can produce gain changes similar to those achieved by conventional intrasaccadic-step adaptation, the changes brought about by reinforcement are substantially slower. One can imagine several reasons for this difference. First, regardless of the mechanism used, saccade adaptation serves to orient the eyes toward a target. However, in our “purest” reinforcement paradigm, auditory reinforcement alone, there is no target at all after the saccade. This condition might be pushing the boundaries of saccade adaptation too far for rapid adaptation. Second, if normal saccade adaptation uses all three mechanisms (retinal error, prediction error, and reinforcement) simultaneously, it may be that the reinforcement component is the component with the longest time constant. [The presence of fast and slow components of saccade adaptation is well established (Chen-

Fig. 8. *A*: time course of the change in saccadic gain in *experiment 5* (subjects *S12*, *S13*, *S14*, and *S15*). Solid circles, mean saccadic gain (averaged for each trial across subjects) in the reinforced decreased gain adaptation sessions; shaded circles, mean saccadic gain (averaged for each trial across subjects) in the conventional decreased gain adaptation sessions. The vertical gray line indicates the beginning of adaptation. Black curve, fits for the reinforced adaptation; gray curve, fits for the conventional paradigm. *B*: time course of the change in saccadic gain in *experiments 1, 3, and 4* (black line and solid circles) and downward conventional paradigm (subjects *S3*, *S16*, and *S17*; gray curve and shaded circles). Symbols are as in *A*. *C*: main sequence parameter *a* during adaptation as a function of main sequence parameter *a* during unreinforced baseline in *experiment 5* (subjects *S12*, *S13*, *S14*, and *S15*). Open symbols, reinforced adaptation; solid symbols, conventional adaptation. *D*: main sequence parameter *a* during adaptation as a function of main sequence parameter *a* during unreinforced baseline in *experiments 1–4* and conventional paradigms (subjects *S3*, *S16*, *S17*, *S18*, *S19*, and *S20*). *E*: mean latencies during adaptation as a function of mean latency during unreinforced baseline in *experiment 5*. Symbols are as in *C*. *F*: mean latencies during adaptation as a function of mean latency during unreinforced baseline in *experiments 1–4* and conventional paradigms (subjects *S3*, *S16*, *S17*, *S18*, *S19*, and *S20*). Symbols are as in *D*.



Harris et al. 2008; Kojima et al. 2004)]. Third, in the case of *experiment 3*, in which a postsaccadic target view is used as reinforcement, the fact that the decreased gain adaptation is faster than the increased gain recovery might be explained by the fact that in the decreased gain situation, the target has made a back step, which might involve the prediction error mechanism, whereas during the reinforced recovery, the target back step would counter the reinforcement contingency.

Beyond these considerations, we must consider the possibility that our using operant conditioning may have accomplished a change in saccade amplitude by a means other than those used by saccade adaptation, that is, we may have turned our subjects into actors, working for a reward by simulating saccade adaptation. We are not inclined to this view for several reasons. First, because our targets appeared at a fixed eccentricity and direction from a constant fixation point on the screen, it would seem that the easiest way to obtain the reward would be to make saccades to a particular place on the screen by making saccades of the appropriate amplitude. One might expect that a subject using a conscious strategy to maximize reward would figure out the most rewarding saccades to make in fewer than 1,000 trials over several days. Second, if subjects were using such a strategy, it would seem curious that the adaptation transferred almost perfectly to unreinforced saccades to targets appearing at different distances from different fixation points (97% and 98% transfer for the position and

amplitude transfer trials), transfer that is similar to that shown by saccades adapted conventionally (Fig. 5, *E* and *F*). Third, when we had the same subject do both reinforced adaptation and conventional adaptation, the degree of adaptation we observed with reinforcement (24% decrease) was comparable to that obtained with conventional adaptation (25% decrease, comparable final gains of 0.67 and 0.70, respectively; Fig. 6), although the reinforcement paradigms involved more trials, and the resulting changes in gain were more variable. Finally, if the subjects were simulating adaptation, one might expect that when put in a conventional adaptation recovery situation, which involved making a saccade to a visible unmoving target, as in *experiment 5*, that gain change would occur extremely quickly. However, as previously discussed, recovery in this situation is gradual and quite similar after reinforced and conventional adaptation (Fig. 5, *C* and *D*), further pointing to similarities between the two types of adaptation.

As mentioned in the RESULTS, some subjects adapted rapidly in the reinforced sessions, as rapidly as did subjects in the conventional adaptation sessions, whereas other subjects adapted much more slowly. To compare the time course of reinforcement-guided and conventional adaptation in the same subjects, we averaged trials across experimental runs for the two types of gain-decreasing adaptations in *experiment 5* and fitted them with second-degree exponentials (Fig. 8*A*) across all adaptation sessions. We estimated the first (i.e., high-frequency) rate constant to be 183 trials for the reinforcement-

guided adaptation versus 49 trials for the conventional paradigm. However, when we compared the data obtained in *experiments 1, 3, and 4* with the data obtained with the conventional adaptation paradigm (*subjects S3, S16, and S17*), the two curves were quite similar and reached similar asymptotes, with slightly different rates (108 vs. 86, respectively, for the first rate constant; Fig. 8B). These values reflect the initial change in gain at the start of adaptation, rather than the final gain achieved. One can speculate that the rates of adaptation may depend on the differing value of the rewards, even though similar underlying processes may be at play. Curiously, the rate of conventional adaptation in the subjects in *experiment 5* who had previously had the reinforced adaptation was faster than in the other subjects (49 vs. 86), perhaps because of some facilitation of adaptation as a result of their previous reinforced adaptation (Ethier et al. 2008; Kojima et al. 2004).

To evaluate whether the saccade themselves differ between the reinforced and conventional adaptation, we compared the dynamics of saccades by plotting the peak velocity of each movement with respect to the saccade amplitude using the following equation (Lebedev et al. 1996):

$$a = \frac{\text{Peak velocity}}{\sqrt{\text{Movement amplitude}}}$$

Parameter a , which is directly proportional to peak velocity, was not different between the reinforced and conventional conditions in the last adaptation sessions in the same subjects (paired t -test, $P = 0.15$, *experiment 5*). However, in general, parameter a slightly increased from the baseline to the adaptation sessions in the case of reinforced, but not conventional, adaptation. This is evident in *experiment 5*, in which the same subjects were adapted under both paradigms (for the last reinforced adaptation sessions, $a = 125$ compared with 108 for the baseline session, paired t -test, $P = 0.02$; for the last conventional adaptation sessions, $a = 114$ compared with 112 for the baseline sessions, paired t -test, $P = 0.11$). It is also evident in the other reinforcement adaptation experiments ($a = 112$ vs. 105 for *experiments 1–4* vs. their baseline sessions, paired t -test, $P = 0.03$; Fig. 8D), but not in the decreased gain conventional adaptation experiments ($a = 105$ for both the adaptation and baseline session, paired t -test, $P = 0.97$). This small difference does not clearly resolve whether the saccades are made by different mechanisms because other studies have shown that reinforcing a specific subset of saccades causes the peak velocities of these movements to increase slightly (Takikawa et al. 2002; Xu-Wilson et al. 2009).

Finally, the latencies of saccades adapted by reinforcement versus conventional paradigms were quite similar (210 vs. 190 ms for the last adaptation sessions of *experiment 5*, t -test, $P = 0.25$). There was a small increase in latency from the baseline to the reinforced adaptation sessions (from 183 to 210 ms in *experiment 5*, paired t -test, $P = 0.39$, Fig. 8E; from 197 to 227 ms in *experiments 1–4*, paired t -test, $P = 0.02$, Fig. 8F), whereas with conventional adaptation, the latencies remained mostly unchanged (190 compared with 178 ms in *experiment 5*, paired t -test, $P = 0.22$, Fig. 8E; 191 compared with 189 ms in the other subjects who experienced conventional adaptation, paired t -test, $P > 0.47$, Fig. 8F).

Significance of Control of Saccadic Gain as a Reinforcement Process

Two established features of saccadic adaptation are compatible with reinforcement being involved. First, the timing of the postsaccadic visual error affects the amount of adaptation: if the target is extinguished during the saccade and a delay is introduced between the end of the saccade and the adaptation step, the amount of adaptation gradually decreases as a function of the delay in both monkeys (Shafer et al. 2000) and humans (Bahcall and Kowler 2000; Fujita et al. 2002). This effect might be interpreted as the necessary response-reinforcer temporal contiguity (Donahoe and Palmer 1994). Second, the sensitivity of saccadic adaptation to contexts such as eye position (Alahyane and Pelisson 2004), head position (Shelhamer and Clendaniel 2002), target distance (Chaturvedi and van Gisbergen 1997), or visual properties of the target (Herman et al. 2009) could also be explained in terms of operant conditioning: a stimulus that signals a probability of reinforcement becomes a discriminative stimulus, thereby influencing the probability of emitting the operant response.

The present experiments provide strong evidence that saccadic gain can depend on learned contingencies. Other attributes of saccades have been shown to be sensitive to reinforcement (Madelain et al. 2011), such as peak velocity (Takikawa et al. 2002) and latency (Ikeda and Hikosaka 2003; Lauwereyns et al. 2002; Madelain et al. 2007; Watanabe et al. 2003b). These results and those of the present study suggest that the oculomotor system might use reinforcement learning to adapt to sensorimotor changes.

GRANTS

This work was supported by National Institutes of Health Grants 1-R01-EY-019508 (to L. Madelain and J. Wallman) and RR-03060 (to J. Wallman), Agence Nationale de la Recherche Grant ANR-09-JCJC-0134 (to L. Madelain and C. Paeye), and a Fulbright Fellowship (to L. Madelain).

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the author(s).

REFERENCES

- Abel LA, Schmidt D, Dell'Osso LF, Daroff RB. Saccadic system plasticity in humans. *Ann Neurol* 4: 313–318, 1978.
- Alahyane N, Pelisson D. Eye position specificity of saccadic adaptation. *Invest Ophthalmol Vis Sci* 45: 123–130, 2004.
- Alahyane N, Pelisson D. Long-lasting modifications of saccadic eye movements following adaptation induced in the double-step target paradigm. *Learn Mem* 12: 433–443, 2005.
- Bahcall DO, Kowler E. The control of saccadic adaptation: implications for the scanning of natural visual scenes. *Vision Res* 40: 2779–2796, 2000.
- Bayer HM, Glimcher PW. Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron* 47: 129–141, 2005.
- Becker W. The neurobiology of saccadic eye movements. *Metrics Rev Oculomot Res* 3: 13–67, 1989.
- Brainard DH. The psychophysics toolbox. *Spat Vis* 10: 433–436, 1997.
- Chaturvedi V, van Gisbergen JA. Specificity of saccadic adaptation in three-dimensional space. *Vision Res* 37: 1367–1382, 1997.
- Chen-Harris H, Joiner WM, Ethier V, Zee DS, Shadmehr R. Adaptive control of saccades via internal feedback. *J Neurosci* 28: 2804–2813, 2008.
- Darceville JC, Madelain L, Buquet C, Charlier J, Miossec Y. Operant conditioning of the visual smooth pursuit in young infants. *Behav Processes* 46: 131–139, 1999.
- Deubel H. Separate adaptive mechanisms for the control of reactive, volitional, and memory guided saccadic eye movements. In: *Attention and*

- Performance. XVII. Cognitive Regulation of Performance: Interaction of Theory and Application, edited by Gopher D, Koriati A. Cambridge, MA: MIT Press, 1999, p. 697–721.
- Donahoe JW, Burgos JE, Palmer DC. A selectionist approach to reinforcement. *J Exp Anal Behav* 60: 17–40, 1993.
- Donahoe JW, Palmer DC. *Learning and Complex Behavior*. Boston, MA: Allyn & Bacon, 1994.
- Ethier V, Zee DS, Shadmehr R. Spontaneous recovery of motor memory during saccade adaptation. *J Neurophysiol* 99: 2577–2583, 2008.
- Fujita M, Amagai A, Minakawa F, Aoki M. Selective and delay adaptation of human saccades. *Brain Res Cogn Brain Res* 13: 41–52, 2002.
- Galbicka G. Shaping in the 21st century: moving percentile schedules into applied settings. *J Appl Behav Anal* 27: 739–760, 1994.
- Galbicka G, Kautz MA, Jagers T. Response acquisition under targeted percentile schedules: a continuing quandary for molar models of operant behavior. *J Exp Anal Behav* 60: 171–184, 1993.
- Harris CM. Does saccadic undershoot minimize saccadic flight-time? A Monte-Carlo study. *Vision Res* 35: 691–701, 1995.
- Harwood MR, Madelain L, Krauzlis RJ, Wallman J. The spatial scale of attention strongly modulates saccade latencies. *J Neurophysiol* 99: 1743–1757, 2008.
- Havermann K, Lappe M. The influence of the consistency of postsaccadic visual errors on saccadic adaptation. *J Neurophysiol* 103: 3302–3310, 2010.
- Henson DB. Corrective saccades: effects of altering visual feedback. *Vision Res* 18: 63–67, 1978.
- Herman JP, Harwood MR, Wallman J. Saccade adaptation specific to visual context. *J Neurophysiol* 101: 1713–1721, 2009.
- Hernandez TD, Levitan CA, Banks MS, Schor CM. How does saccade adaptation affect visual perception? *J Vis* 8: 3 11713–16, 2008.
- Hikosaka O, Nakamura K, Nakahara H. Basal ganglia orient eyes to reward. *J Neurophysiol* 95: 567–584, 2006.
- Hopp JJ, Fuchs AF. The characteristics and neuronal substrate of saccadic eye movement plasticity. *Prog Neurobiol* 72: 27–53, 2004.
- Hopp JJ, Fuchs AF. Identifying sites of saccade amplitude plasticity in humans: transfer of adaptation between different types of saccade. *Exp Brain Res* 202: 129–145, 2010.
- Ikeda T, Hikosaka O. Reward-dependent gain and bias of visual responses in primate superior colliculus. *Neuron* 39: 693–700, 2003.
- Kojima Y, Iwamoto Y, Yoshida K. Memory of learning facilitates saccadic adaptation in the monkey. *J Neurosci* 24: 7531–7539, 2004.
- Kommerell G, Olivier D, Theopold H. Adaptive programming of phasic and tonic components in saccadic eye movements. Investigations of patients with abducens palsy. *Invest Ophthalmol Vis Sci* 15: 657–660, 1976.
- Kording KP, Tenenbaum JB, Shadmehr R. The dynamics of memory as a consequence of optimal adaptation to a changing body. *Nat Neurosci* 10: 779–786, 2007.
- Lauwereyns J, Watanabe K, Coe B, Hikosaka O. A neural correlate of response bias in monkey caudate nucleus. *Nature* 418: 413–417, 2002.
- Lebedev S, Van Gelder P, Tsui WH. Square-root relations between main saccadic parameters. *Invest Ophthalmol Vis Sci* 37: 2750–2758, 1996.
- Madelain L, Champrenaut L, Chauvin A. Control of sensorimotor variability by consequences. *J Neurophysiol* 98: 2255–2265, 2007.
- Madelain L, Harwood MR, Herman JP, Wallman J. Saccade adaptation is unhampered by distractors. *J Vision* 10: 29, 2010.
- Madelain L, Krauzlis RJ. Effects of learning on smooth pursuit during transient disappearance of a visual target. *J Neurophysiol* 90: 972–982, 2003.
- Madelain L, Krauzlis RJ, Wallman J. Spatial deployment of attention influences both saccadic and pursuit tracking. *Vision Res* 45: 2685–2703, 2005.
- Madelain L, Paeye C, Darcheville JC. Operant control of human eye movements. *Behav Processes* 87: 142–148, 2011.
- McLaughlin SC. Parametric adjustment in saccadic eye movements. *Percept Psychophys* 2: 359–362, 1967.
- Montagnini A, Chelazzi L. The urgency to look: prompt saccades to the benefit of perception. *Vision Res* 45: 3391–3401, 2005.
- Nakamura K, Hikosaka O. Role of dopamine in the primate caudate nucleus in reward modulation of saccades. *J Neurosci* 26: 5360–5369, 2006.
- Noto CT, Robinson FR. Visual error is the stimulus for saccade gain adaptation. *Brain Res Cogn Brain Res* 12: 301–305, 2001.
- Noto CT, Watanabe S, Fuchs AF. Characteristics of simian adaptation fields produced by behavioral changes in saccade size and direction. *J Neurophysiol* 81: 2798–2813, 1999.
- Paeye C, Madelain L. Reinforcing saccadic amplitude variability. *J Exp Anal Behav* 95: 149–162, 2011.
- Pelisson D, Alahyane N, Panouilleres M, Tilikete C. Sensorimotor adaptation of saccadic eye movements. *Neurosci Biobehav Rev* 34: 1103–1120, 2010.
- Pelli DG. The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spat Vis* 10: 437–442, 1997.
- Robinson FR, Noto CT, Bevans SE. Effect of visual error size on saccade adaptation in monkey. *J Neurophysiol* 90: 1235–1244, 2003.
- Sudder CA, Batourina EY, Tunder GS. Comparison of two methods of producing adaptation of saccade size and implications for the site of plasticity. *J Neurophysiol* 79: 704–715, 1998.
- Semmlow JL, Gauthier GM, Vercher JL. Mechanisms of short-term saccadic adaptation. *J Exp Psychol Hum Percept Perform* 15: 249–258, 1989.
- Shafer JL, Noto CT, Fuchs AF. Temporal characteristics of error signals driving saccadic gain adaptation in the macaque monkey. *J Neurophysiol* 84: 88–95, 2000.
- Shelhamer M, Clendaniel RA. Context-specific adaptation of saccade gain. *Exp Brain Res* 146: 441–450, 2002.
- Skinner BF. Selection by consequences. *Science* 213: 501–504, 1981.
- Straube A, Fuchs AF, Usher S, Robinson FR. Characteristics of saccadic gain adaptation in rhesus macaques. *J Neurophysiol* 77: 874–895, 1997.
- Sugrue LP, Corrado GS, Newsome WT. Matching behavior and the representation of value in the parietal cortex. *Science* 304: 1782–1787, 2004.
- Takikawa Y, Kawagoe R, Itoh H, Nakahara H, Hikosaka O. Modulation of saccadic eye movements by predicted reward outcome. *Exp Brain Res* 142: 284–291, 2002.
- Wallman J, Fuchs AF. Saccadic gain modification: visual error drives motor adaptation. *J Neurophysiol* 80: 2405–2416, 1998.
- Watanabe K, Lauwereyns J, Hikosaka O. Effects of motivational conflicts on visually elicited saccades in monkeys. *Exp Brain Res* 152: 361–367, 2003a.
- Watanabe K, Lauwereyns J, Hikosaka O. Neural correlates of rewarded and unrewarded eye movements in the primate caudate nucleus. *J Neurosci* 23: 10052–10057, 2003b.
- Wolpert DM, Ghahramani Z, Jordan MI. An internal model for sensorimotor integration. *Science* 269: 1880–1882, 1995.
- Wong AL, Shelhamer M. Sensorimotor adaptation error signals are derived from realistic predictions of movement outcomes. *J Neurophysiol* 105: 1130–1140, 2011.
- Xu-Wilson M, Zee DS, Shadmehr R. The intrinsic value of visual information affects saccade velocities. *Exp Brain Res* 196: 475–481, 2009.

REINFORCING SACCADIC AMPLITUDE VARIABILITY

CÉLINE PAEYE AND LAURENT MADELAIN

UNIVERSITÉ CHARLES DE GAULLE LILLE III

Saccadic endpoint variability is often viewed as the outcome of neural noise occurring during sensorimotor processing. However, part of this variability might result from operant learning. We tested this hypothesis by reinforcing dispersions of saccadic amplitude distributions, while maintaining constant their medians. In a first experiment we reinforced the least frequent saccadic amplitudes to increase variability, and then reinforced the central part of the amplitude distributions to reduce variability. The target was placed at a constant distance from the fovea after the saccade to maintain the postsaccadic visual signal constant and an auditory reinforcement was delivered depending on saccadic amplitude. The second experiment tested the effects of the contingency. We reinforced high levels of variability in 4 participants, whereas 4 other participants were assigned to a yoked control group. On average, saccadic amplitude standard deviations were doubled while the medians remained mostly unchanged in the experimental participants in both experiments, and variability returned to baseline level when low variability was reinforced. In the control group no consistent changes in amplitude distributions were observed. These results, showing that variability can be reinforced, challenge the idea of a stochastic neural noise. We instead propose that selection processes constrain saccadic amplitude distributions.

Key words: variability, operant, motor control, shaping, selection, ocular saccade, humans

Ocular saccades are the rapid movements of the eyes that place the retinal image of a target on the fovea (the small high-acuity area of the retina), allowing the visual system to perceive the fine details of the visual environment. Because of the limited span of the fovea, saccades require a fine motor control as well as adaptive properties to maintain accuracy in response to changes occurring in the saccadic system or in the environment. Saccadic amplitude adaptation is a compelling example of the adaptive properties of the primate visuomotor system to changes occurring during the lifespan—such as growing, aging or ocular pathologies (Abel, Schmidt, Dell’Osso, & Daroff, 1978; Kommerell, Olivier, & Theopold, 1976). Saccadic amplitude is the distance traveled by the eye between two fixation points: When making a saccade toward an

intended target the amplitude is usually such that the eyes land close to the object. However, pathological saccadic dysmetria is sometimes observed, as in patients suffering from monocular muscular weakness in whom saccades fall short of the intended goal. Amplitude adaptation may then be induced by placing a patch on the nonaffected eye to block vision: Saccadic amplitude progressively recovers (Optican, Zee, & Chu, 1985; Zee, Optican, Cook, Robinson, & Engel, 1976). Saccadic adaptation has also been demonstrated in the laboratory using an intrasaccadic step paradigm (McLaughlin, 1967) consisting in having the target surreptitiously jump backward or forward during the saccade so that the gaze lands away from the target. This induced mismatch between the target displacement and the saccadic amplitude is progressively compensated by the oculomotor system and the eye lands closer and closer to the target’s position (see Hopp & Fuchs, 2004; Pelisson, Alahyane, Panouilleres, & Tilikete, 2010, for reviews).

It has been proposed that the postsaccadic error—the retinal eccentricity of the target after the saccade lands—is the feedback used to correct subsequent eye movements, leading to changes in saccadic gain—the ratio of the saccadic amplitude to the displacement of the target (Noto & Robinson, 2001; Robinson,

This research was supported in part by National Institute of Health Grant 1R01EY019508 (LM) and Agence Nationale Pour la Recherche Grant ANR-JC09_494068. Portions of these data were presented at the 2010 Society for the Quantitative Analyses of Behavior convention, San Antonio, TX. We thank Allen Neuringer and an anonymous reviewer for their helpful comments on an earlier version of the manuscript.

Address correspondence to the second author at Laboratoire Ureca, PRES Lille Nord de France, Université de Lille III, Domaine universitaire du Pont de Bois, BP 149 59653 Villeneuve d’Ascq Cedex, France (e-mail: laurent.madelain@univ-lille3.fr).

doi: 10.1901/jeab.2011.95-149

Noto, & Bevens, 2003; Robinson, Soetedjo, & Noto, 2006; Seeberger, Noto, & Robinson, 2002; Wallman & Fuchs, 1998). According to Bahcall and Kowler (2000), the oculomotor system might use elaborated comparisons between the actual postsaccadic retinal error and a predicted retinal image. In a similar vein, computational motor control theories postulate a forward model that predicts the outcome of the movement, the result of the comparison with the actual retinal error being used to adjust the motor command when necessary (Chen-Harris, Joiner, Ethier, Zee, & Shadmehr, 2008; Wolpert & Ghahramani, 2000). Saccadic adaptation would therefore imply a servo-mechanism in which the magnitude of the output is capable of controlling the input by way of a hypothetical internal comparator, a proposition that is difficult to probe at the behavioral level (see Ingvaldsen & Whiting, 1997, for a related discussion).

Contrasting with these computational hypotheses, there is support for the idea that saccades have attributes of operant behavior. For instance, when monkeys had to perform a saccade in a direction associated with food, saccade peak velocities were higher, trajectories straighter and latencies shorter than in a nonreinforced direction (Lauwereyns, Watanabe, Coe, & Hikosaka, 2002; Takikawa, Kawagoe, Itoh, Nakahara, & Hikosaka, 2002; Watanabe, Lauwereyns, & Hikosaka, 2003). Activity changes were found in brain structures such as the superior colliculus (Ikeda & Hikosaka, 2003), the substantia nigra pars reticulata (Sato & Hikosaka, 2002) and the caudate nucleus (Lauwereyns et al., 2002; Nakamura & Hikosaka, 2006; Watanabe et al., 2003) depending on the reinforcement condition, a feature shared by other operant behaviors (Schultz, 2000). In humans, Xu-Wilson, Zee, and Shadmehr (2009) found that saccades had higher peak velocities and shorter durations when they were made to a stimulus with a high value (faces) than to a stimulus with a low value (random pixels). Montagnini and Chelazzi (2005) manipulated the difficulty of a letter discrimination task by reducing the delay between saccade initiation and target onset, so that the ability to perform the task (that is, to clearly see the target) required the subjects to reduce their saccade latency. Under these conditions saccadic latencies decreased and peak velocities in-

creased. Moreover, we recently found that saccadic reaction time variability might be controlled by reinforcement (Madelain, Champrenaut, & Chauvin, 2007), further demonstrating that saccadic eye movements are operant behaviors. Interestingly, there is evidence that visual pursuit—another kind of voluntary eye movements—might also be controlled by learned contingencies (Darcheville, Madelain, Buquet, Charlier, & Miossec, 1999; Madelain & Krauzlis, 2003).

A critical feature of operant behaviors is that their variability might be controlled by reinforcement (Neuringer, 2002). A large body of research has been devoted to studying operant variability in various dimensions of behaviors, such as interresponse times (Blough, 1966), topography (Goetz & Baer, 1973; Pryor, Haag, & O'Reilly, 1969; Stokes, 1995) or response sequences in animals (e.g. Abreu-Rodrigues, Lattal, dos Santos, & Matos, 2005; Machado, 1989, 1992, 1997; Neuringer, 1992; Page & Neuringer, 1985) as well as in humans (e.g. Miller & Neuringer, 2000; Neuringer, 1986; Stokes & Balsam, 2001; see Lee, Sturmey, & Fields, 2007, and Neuringer, 2002, for reviews). We propose that saccadic endpoint variability may be placed under operant control as well. This hypothesis contrasts with the conventional view stating that sensorimotor variability results from an uncontrollable stochastic neural noise that affects each stage between a sensory event and the motor response—sensing, information processing, movement planning and executing (Van Beers, 2007; see Faisal, Selen, & Wolpert, 2008, for a review).

One could hypothesize that, in real life, specific levels of tolerated—or required—saccadic endpoint variability may be reinforced by the consequences of a movement, namely the clear vision of targets. For instance, a low level of variability is helpful when reading: Given that letters in peripheral vision appear too blurred to be distinguished, eye movements must accurately orient the fovea to specific locations in order to see each word of a sentence. On the other hand saccade accuracy is not relevant when facing a blank page: Eye movement consequences are the same whichever location is targeted and here the constraints acting on endpoint variability may be looser such that any level of variability would be equally efficient. To test the hypoth-

esis that part of saccadic variability may result from operant learning, we manipulated the reinforcement contingencies of saccadic amplitude variability and induced changes in the spread of the amplitude distributions.

EXPERIMENT 1: EXTENT OF CONTROL BY REINFORCEMENT CONTINGENCIES

METHOD

Rationale and Subjects

This experiment aimed at probing the extent of control one can exert on saccadic amplitude variability. The two authors and an undergraduate student (with normal or corrected to normal vision) had first to increase the spread of their amplitude distribution while maintaining the median constant, and then a comparable decrease in variability was required.

Auditory reinforcement was delivered on a trial-to-trial basis depending on the saccadic amplitude, the criteria being continuously updated. It was necessary to control the postsaccadic retinal error so that it remained constant after each saccade and matched the individual baseline saccadic gain (i.e., the ratio of the saccadic amplitude to the target displacement). Four experimental conditions were performed. All experimental procedures were reviewed and approved by the Institutional Review Board and each participant gave informed consent.

Regular Baseline

Four sessions of 200 regular saccades were used to assess the baseline saccadic gain in each participant, that is, the median gain computed over these baseline saccades. Each trial started with a random fixation period of 500–1000 ms during which a white fixation cross was displayed pseudorandomly between 2.4° and 13° to the left or right of the center of the screen against a grey background (luminance 40 cd/m²). When the fixation cross was extinguished the target stimulus, a luminance-defined Gaussian patch ($SD=0.93^\circ$, maximum luminance 62 cd/m²), appeared between 9.5° and 14.2° horizontally from the fixation cross. Subjects were required to make a saccade to this target within 700 ms following target appearance. If the latency was less than 100 ms or longer than 700 ms the trial was canceled and the fixation cross immediately

reappeared in place of the target. In a trial, the target was displayed for a total duration of 1200 ms. We usually recorded two sessions per day; participants had about 10 min break between sessions, during which they were free to move.

Stabilized Baseline

For four other sessions of 200 trials, as well as for the subsequent experimental conditions, we stabilized the target image on the fovea at a distance depending on the individual baseline saccadic gain. That is, when the saccade was detected the target was extinguished and it reappeared at a new location relative to the eye position at the end of the movement. This location was obtained by multiplying the landing point of the current saccade (e.g. 12° to the right of the fixation point) by the reciprocal of the baseline gain (e.g., 0.9) obtained over the regular baseline. Therefore, for a participant in whom the baseline saccadic gain was less than 1 the target appeared slightly further from the postsaccadic eye position. In our example the target would be displayed $(1/0.9)*12 = 13.33^\circ$ deg to the right of the fixation point, as illustrated in Figure 1A. This manipulation was designed to avoid the undesired change of gain that might be induced when the target image is stabilized on the fovea without this correction, because of the tendency of the saccadic system to maintain its gain (Havermann & Lappe, 2010; Henson, 1979; Robinson et al., 2003). A new trial began with the reappearance of the fixation cross. The median gain measured during these 800 stabilized trials was then used as the goal gain in the subsequent experimental conditions.

Learning

Increases in saccadic amplitude variability were induced by reinforcing least frequent amplitudes: Saccades of rare amplitudes were reinforced but saccades of frequent amplitudes were not. Reinforcement criteria were computed using 10 gain bins centered on the individual goal gain for each participant. Because of the common tendency of having a saccadic gain slightly less than 1 we used unequal bins: The size of the five lower bins were 0.06 units of gain whereas the size of the five upper bins were 0.03 units of gain.

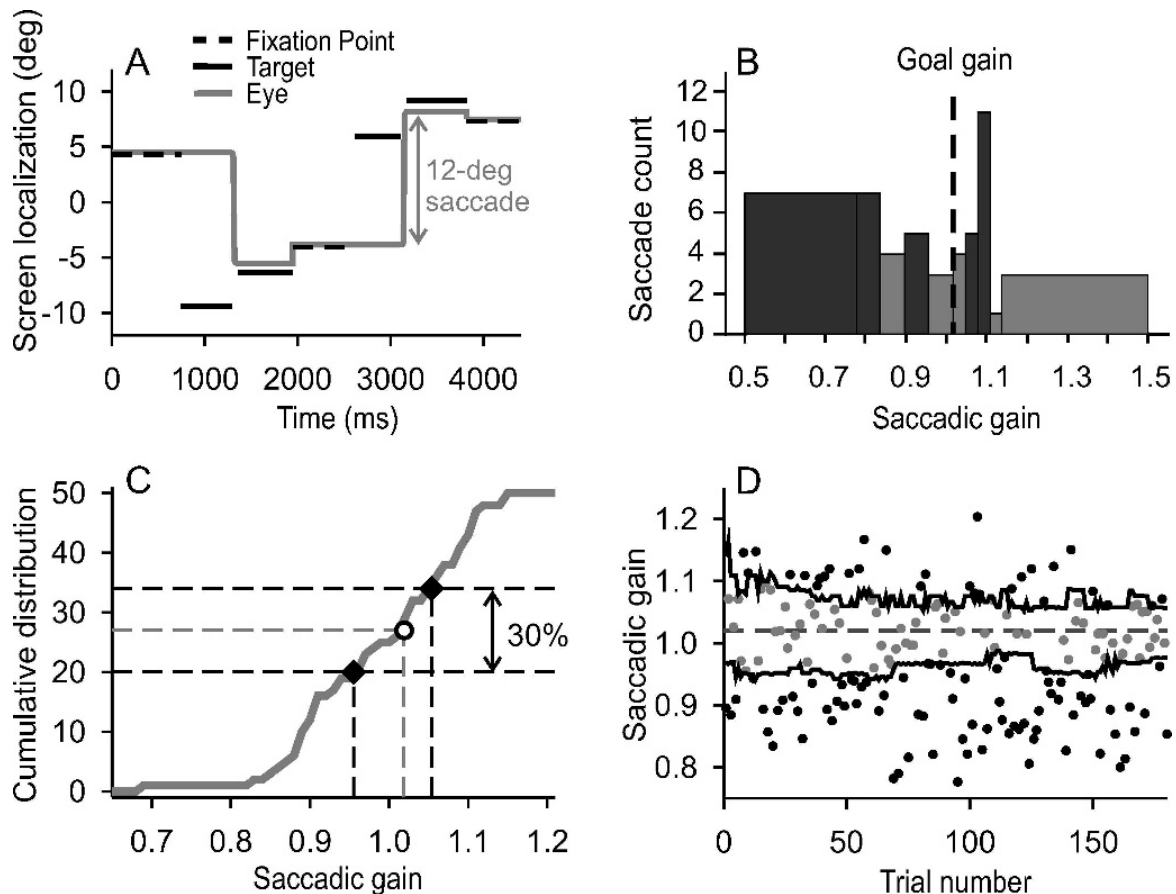


Fig. 1. A: Representative examples of fixation point, target, and eye positions for two trials during stabilized baseline (participant S3, baseline saccadic median gain 0.92). B: Gain bins and reinforcement criteria for a single learning trial (participant S2). A tone was emitted if the current gain fell into one of the five least frequent bins (light grey). C: Computation of reinforcement criteria for a single recovery trial (participant S2). Solid grey line, cumulative gain distribution for the previous 50 trials. White circle, goal gain. Black diamonds, saccadic gain values $\pm 15\%$ around the goal gain; vertical black dashed lines, reinforcement criteria. D: Example of a recovery session (participant S2): changes in reinforcement criteria (solid black lines), reinforced saccades (grey points) and nonreinforced saccades (black points). Gray dashed line represents goal median gain.

Figure 1B shows an example for a single learning trial of participant S2 whose goal gain was 1.019.

The lowest and highest bins values were set to 0.5 and 1.5. For instance, for participant S2 the bin intervals were $[0.5; 0.779]$, $[0.779; 0.839]$ and so on, the last ones being $[1.079; 1.109]$ and $[1.109; 1.5]$. For each trial we computed the amplitude frequencies for each bin over the previous 50 saccades. Each new value then replaced the oldest one so that the frequencies were always computed with respect to the 50 most recent trials. For the first trial of a session the frequencies were obtained over 50 trials randomly chosen from the previous session. If the current gain fell into one of the five least frequent bins a tone (a 60 ms 1000 Hz sine wave) was emitted.

Furthermore, whenever three out of five consecutive saccades (including the fifth one) were followed by the tone, another auditory reinforcer was delivered (a synthetic voice saying “bravo”) indicating that the subject won one point. The number of learning sessions (200 trials) depended on the time needed to double the standard deviations of the amplitude distributions with respect to the stabilized baseline standard deviation (S1, S2 and S3 performed 40, 18 and 32 learning sessions, respectively).

Recovery

During this fourth experimental condition the reinforcement contingencies aimed at decreasing the level of amplitude variability. The differential reinforcement criteria were

systematically changing from trial to trial but they were now computed based on a targeted percentile reinforcement procedure (Galbicka, 1994; Galbicka, Kautz, & Jagers, 1993), similar to the ones we used to shape the distributions of saccadic latencies (Madelain et al., 2007) and to alter saccadic gain (Madelain, Paeye, & Wallman, 2008).

This procedure involved two simultaneous criteria which control the range of reinforced amplitudes below and above the goal gain. Specifically, the cumulative distribution of the 50 most recent saccadic amplitudes was used to compute the range of reinforced saccades: The lower and upper reinforcement criteria corresponded to the gains of the seventh saccades (15% of the previous 50) which were closest (below and above) to the goal gain (the median gain obtained in the stabilized baseline trials). By reinforcing only a fraction of saccades ($\pm 15\%$ in our experiments) closest to the goal gain we expected the amplitude distribution to become sharper and sharper and peak near the goal gain. Figure 1C illustrates how the reinforcement criteria were computed for a single trial: The goal gain corresponded to the stabilized baseline gain (1.019); based on the cumulative gain distribution computed over the last 50 trials, the lower criterion was the gain of the seventh saccade smaller than the goal gain (0.955) and the upper criterion corresponded to the gain of the seventh saccade larger than the goal gain (1.054). On the subsequent trial the saccade would be reinforced if its gain fell between 0.955 and 1.054. Each new value replaced the oldest one. For the first 50 trials of a session the cumulative gain distribution was computed based on 50 saccades randomly extracted from the previous session. Figure 1D illustrates the changes in the reinforcement criteria and in the reinforced gains on a trial-to-trial basis during a recovery session. With this procedure every trial might be reinforced if the gain fell within the reinforcement criteria but the overall probability of reinforcement remained approximately constant across trials. As in the learning sessions, participants were encouraged to maximize their local rate of reinforcement by giving one point signaled by an auditory stimulus as soon as three out of five consecutive trials were reinforced. This aimed at constraining the amplitude criteria around the goal gain, thereby having the

variability decrease. The number of recovery sessions (200 trials) necessary to reduce the standard deviations was 18, 22 and 6 for S1, S2 and S3 respectively.

EXPERIMENT 2: IS THE CONTINGENCY BETWEEN EYE MOVEMENTS AND CONSEQUENCES NECESSARY TO CONTROL VARIABILITY? YOKED CONTROL

Subjects

We had 8 naïve participants perform this experiment to ensure that auditory reinforcement contingent on saccades, and not the awareness of the rules or the image stabilization trick, drove the changes in amplitude variability.

The 8 participants had normal or corrected to normal vision. They had no previous experience in oculomotor experiments. To familiarize them with the eye-movement recording apparatus and calibration procedure we first had them make 50 saccades using the same stimuli as in the first regular baseline (we do not report results from these trials). All experimental procedures were reviewed and approved by the Institutional Review Board and each participant gave informed consent.

Procedure

Participants were first instructed to make a saccade to the appearance of the target after fixating the cross. After the two baseline conditions they were told that the tones were emitted depending on the variability of their saccadic amplitude. They were instructed to earn as many points as possible as a game and the winner was rewarded with a prize. Nothing was explained about our expectancy to increase or decrease the variability nor about the target image stabilization.

The design was identical to the one used in Experiment 1 except for a few changes. First, no recovery condition was carried out. Second, during the stabilized baseline trials, saccades were differentially reinforced according to the targeted percentile procedure similar to the one described in Experiment 1, aiming at concentrating saccadic amplitudes around the baseline amplitude. In addition, the fixation cross reappeared pseudorandomly near the target, between $\pm 2.4^\circ$ and $\pm 13^\circ$ from the center of the screen. The median gain

measured during these stabilized trials was then used as the goal gain in the subsequent experimental condition.

Participants were then assigned to two groups of 4. In the experimental group the variability learning procedure lasted until the standard deviations were multiplied by 1.5 with respect to the stabilized baseline standard deviations. The number of learning sessions (200 trials) was 17, 18, 10 and 22 for participants S4 to S7, respectively.

For the yoked control group (S8 to S11) the auditory stimuli were independent of the saccadic amplitude: The tones were programmed in advance and matched the auditory consequences obtained by one of the participants. For example, for S4 the second and fourth trials of the first learning session were reinforced. For the corresponding yoked control participant S8 the second and fourth saccades of the first learning session were followed by the tone regardless of the saccadic amplitudes. Each participant's sequences of reinforcement as well as target presentations were therefore replicated in one of the participants of the yoked control group. The computation of the postsaccadic target position remained linked to the individual baseline saccadic gain. As in the experimental group, trials were aborted when the latency was below 100 ms or above 700 ms. If the instructions, the visual stimuli or the reinforcement intermittency were responsible for increasing amplitude variability, the changes in distributions should be similar in both groups. This would not be the case if the contingency between eye movements and their consequences controlled the saccadic amplitude variability.

Apparatus

Stimuli were generated on a Power Mac G4 using the Psychophysics Toolbox extensions for Matlab (Martinez & Martinez, 2002) and displayed on a video monitor (Iiyama HM204DT, 100 Hz) at a viewing distance of 60 cm. To minimize measurement errors, the subject's head movements were restrained using a dental impression and a forehead rest, so that the eyes in primary gaze position were directed toward the center of the screen. Presentation of stimuli, and acquisition, display, and storage of responses were controlled by a PXI computer using the LabView Real-

Time software package (National Instruments). Trigger signals from the visual display computer to the LabView computer allowed us to synchronize data collection to stimuli presentation with 1-ms resolution. Eye movements were measured continuously with an infrared video-based eye tracking system (ISCAN, RK-726) at 240 Hz. Before each experimental session we calibrated the eye tracker by having subjects repeatedly fixate a set of seven horizontal locations to generate a linear function for converting raw eye tracker values to horizontal eye positions.

Data Acquisition and Analysis

For on-line saccade detection, a real-time algorithm used a point-to-point velocity criterion to identify the start and end of the first saccade within 700 ms after the target onset. To have the target blank at saccade onset and reappear close to the fovea after the saccade, we computed the median eye position over 12 ms following the saccade, and then restored the target at its new location with respect to the fovea as soon as the position computation was completed. On average the target was blanked for 65 ms from saccade onset. On-line measured amplitudes were used to compute the reinforcement criteria.

For off-line analysis of the eye movement recordings, an interactive analysis program was used to display and analyze the data. Horizontal eye velocities and accelerations were obtained by differentiating the eye position signal over an 8 ms window. Saccades were detected by applying a set of fixed velocity and acceleration criteria. A program written in LabView software presented the start and end of each saccade immediately following the target step to the investigator for confirmation or, if necessary, for correction. Saccadic amplitudes were computed by subtracting the median eye positions across 25 ms windows preceding and following the saccade. We measured only the first saccade within 700 ms after the target onset. We discarded trials in which the gain was below 0.5 or above 1.5 or the latency shorter than 100 ms (on average 3.27% of the saccades were discarded).

RESULTS

The saccadic gain distributions varied according to the reinforcement contingencies.

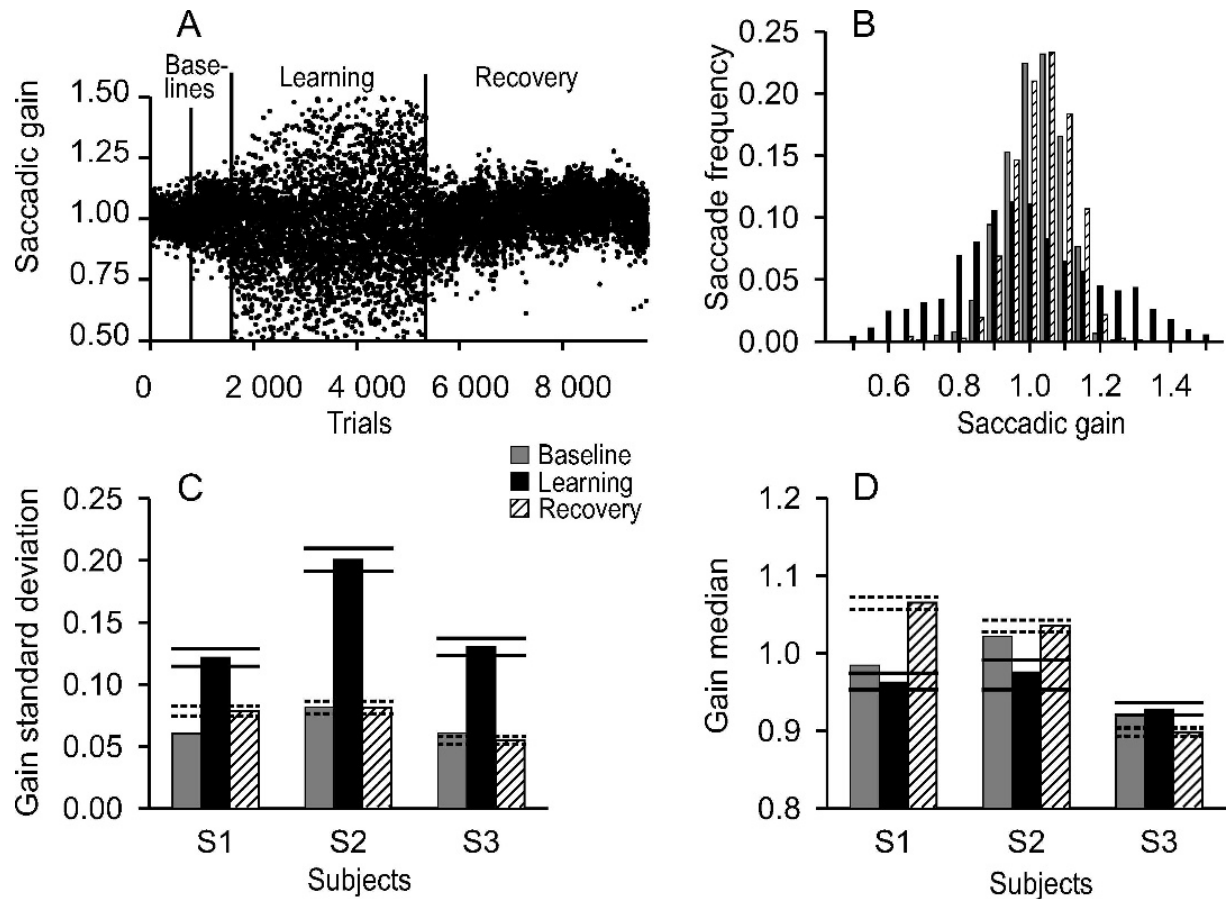


Fig. 2. Results of Experiment 1. A: Saccadic gain in each experimental condition, each data point corresponding to one trial (participant S2). B: Individual frequency distributions (participant S2) of saccadic gains for the stabilized baseline (800 saccades – gray), for the learning trials (last 800 saccades – black) and for the recovery trials (last 800 saccades – hatched). C: Saccadic gain standard deviations for the last four sessions (800 trials) of stabilized baseline (gray), learning (black) and recovery (hatched). 95th bootstrap percentile confidence intervals over the learning and recovery conditions, solid and dashed lines, respectively. D: Corresponding median gains. Bar colors and lines as in panel B.

For all experimental subjects we observed a systematic increase in the dispersion but not in the median gain. In the control group the changes in gain distributions were not consistent across subjects.

Experiment 1

Figure 2A displays the saccadic gain (saccadic amplitude / target displacement amplitude) in each experimental condition for participant S2. Figure 2B summarizes these data by plotting the frequency distributions of saccadic gains for the four last sessions of each experimental condition. The dispersion increased after learning compared to baseline. It was then dramatically reduced at the end of recovery: The frequency distributions of baseline and recovery were almost perfectly super-

imposed. The median gains remained mostly unchanged throughout the three experimental conditions (1.02, 0.98 and 1.04 for the stabilized baseline, the learning condition and the recovery, respectively). The standard deviation changed from 0.08 in the baseline condition to 0.2 at the end of the learning condition and returned to 0.08 at the end of recovery.

To estimate the changes in the standard deviations we computed the 95th bootstrap percentile confidence intervals (Efron, Jolivet, & Hordan, 1995) over the last four learning sessions (Figure 2C). For participant S2 the standard deviations in the baseline and the recovery were clearly outside this confidence interval revealing that the change was statistically significant. This was also true for partic-

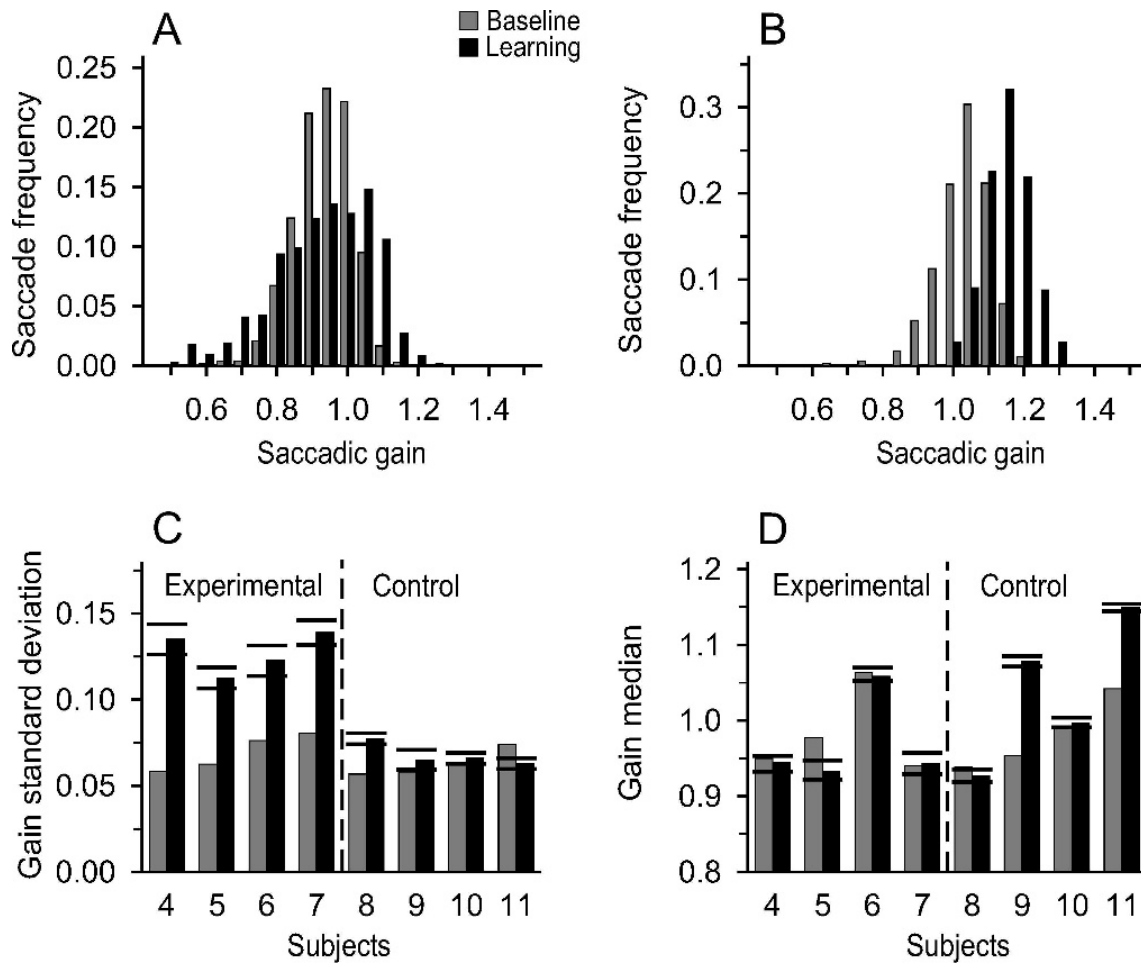


Fig. 3. Results of Experiment 2. A: Individual frequency distributions (experimental participant S7) of saccadic gains for the stabilized baseline (800 saccades – gray) and for the learning trials (last 800 saccades – black). B: Same as A for the control participant S11, yoked with participant S7. C: Saccadic gain standard deviations for the last four sessions (800 trials) of stabilized baseline (gray) and learning (black). 95% bootstrap percentile confidence intervals over the learning condition, solid lines. D: Corresponding median gains. Bar colors and lines as in panel C.

ipants S1 and S3. On average the standard deviations increased by a factor of 2.21 between the baseline and the end of learning. Interestingly, between the baseline and recovery conditions, standard deviations increased by only a factor of 1.06, indicating that the spread of the distribution almost perfectly recovered. The median gains were also altered but the changes were not consistent across participants (Figure 2D). For participants S1 and S2 the medians were significantly lower in the learning sessions whereas for participant S3 it remained unchanged during learning but was slightly reduced at the end of recovery. On average the median changed by a factor of 0.98 from the baseline condition to the learning condition and by a factor of 1.02 from baseline to recovery.

Experiment 2

For the experimental participants the changes in the saccadic gain distributions were analogous to the ones obtained in Experiment 1. As exemplified in Figure 3A for participant S7, the dispersion increased after learning (from 0.08 to 0.14) whereas the median was not altered (0.94 before and at the end of learning). This is not the case for the yoked control participant S11 (figure 3B): The gain standard deviation slightly decreased (from 0.07 to 0.06), but the median increased (from 1.04 to 1.15) after being exposed to the exact same saccades' consequences as participant S7.

The bootstrap percentile confidence intervals presented in Figure 3C reveal that for each experimental participant the gain stan-

dard deviation after learning was significantly higher than the one from baseline, increasing by a factor of 1.86 on average. The control participants also exhibited a change in gain standard deviation but these changes were not consistent: It increased in S8, S9 and S10 but decreased in S11. Moreover these changes were much smaller than those observed in the experimental group: On average standard deviations increased by a factor of 1.1 from the baseline to the end of learning.

Figure 3D plots the median gains as well as the confidence intervals computed over the four last sessions of learning. For 3 out of 4 experimental participants there were no statistical differences while there was a decrease in 1 participant (S5, from 0.98 to 0.93). For the control group the median gain significantly increased in participants S9, S10 and S11 and decreased in participant S8. It should be noted that these changes were large in 2 participants (more than 10% for S9 and S11).

U Values

In addition to the standard deviation, we examined another measure of dispersion, the “uncertainty” or *U* value. This statistic is commonly used to evaluate behavioral variability (e.g. Page & Neuringer, 1985) and does not require assumptions upon the distributions. We computed *U* values using the following formula:

$$U = \frac{-\sum_1^N (p_n \log_2 (p_n))}{\log_2 (N)}$$

where *p* represents the relative frequency of a bin *n* and *N* the number of bins. Here we used 10 equal gain bins of 0.1, from 0.5 to 1.5. *U* value reflects the likelihood that the gain of a trial falls in each bin. If the 10 bins contain equal number of trials, then *U* equals 1. Conversely, if all saccadic amplitudes fall within one single gain bin, *U* equals 0. We computed the *U* values over the last four sessions of each experimental condition. We also computed 95th bootstrap percentile confidence intervals for the *U* values obtained at the end of learning.

Figure 4A (squares and circles) plots the *U* values at the end of learning as a function of the *U* values obtained during baseline in both experiments. All baseline *U* values were close

to each other—ranging from 0.406 to 0.551. The data scattered into two groups: On the one hand, data from the experimental participants (squares and filled circles) are all clearly above equality line, indicating a large increase in the spread of the distribution after learning. On average the *U* values were multiplied by a factor of 1.61 (Experiment 1) and 1.49 (Experiment 2, experimental group). On the other hand, data from the yoked control group (unfilled circles) are close to or even below the equality line, indicating small changes with respect to baseline: *U* values increased by a factor of 1.08 on average after noncontingent reinforcement, but remained unchanged for participant S10 and significantly decreased for S11. Moreover, data from the recovery sessions (down triangles) also plot along the unity line—the *U* values changed only by a factor of 1.05. In 2 out of the 3 subjects the *U* values were lower in the recovery than in the baseline (as confirmed by the 95th bootstrap percentile confidence intervals).

DISCUSSION

This study demonstrates that part of saccadic amplitude variability can be manipulated through learned contingencies. The first experiment shows that reinforcement can induce large changes in saccadic amplitude variability while maintaining the median gains globally unaltered: Compared to stabilized baseline, the standard deviations after learning were multiplied by an average factor of 2.21. Furthermore, at the end of the recovery sessions variability decreased to near baseline level. Therefore, these effects may not be attributed to the target image stabilization.

Instructions could have played a role in increasing variability, as observed elsewhere for interresponse times (Joyce & Chase, 1990), topography (Eisenberger, Armeli, & Pretz, 1998) or response sequences (Hopkinson & Neuringer, 2003; Neuringer, 1986). However, in our second experiment the yoked control participants' data show that instructions alone could not account for the increase in variability because when reinforcement was not contingent on saccadic amplitude variability we did not observe a systematic increase in the distributions spread. Our results extend the findings of Madelain et al. (2007) who

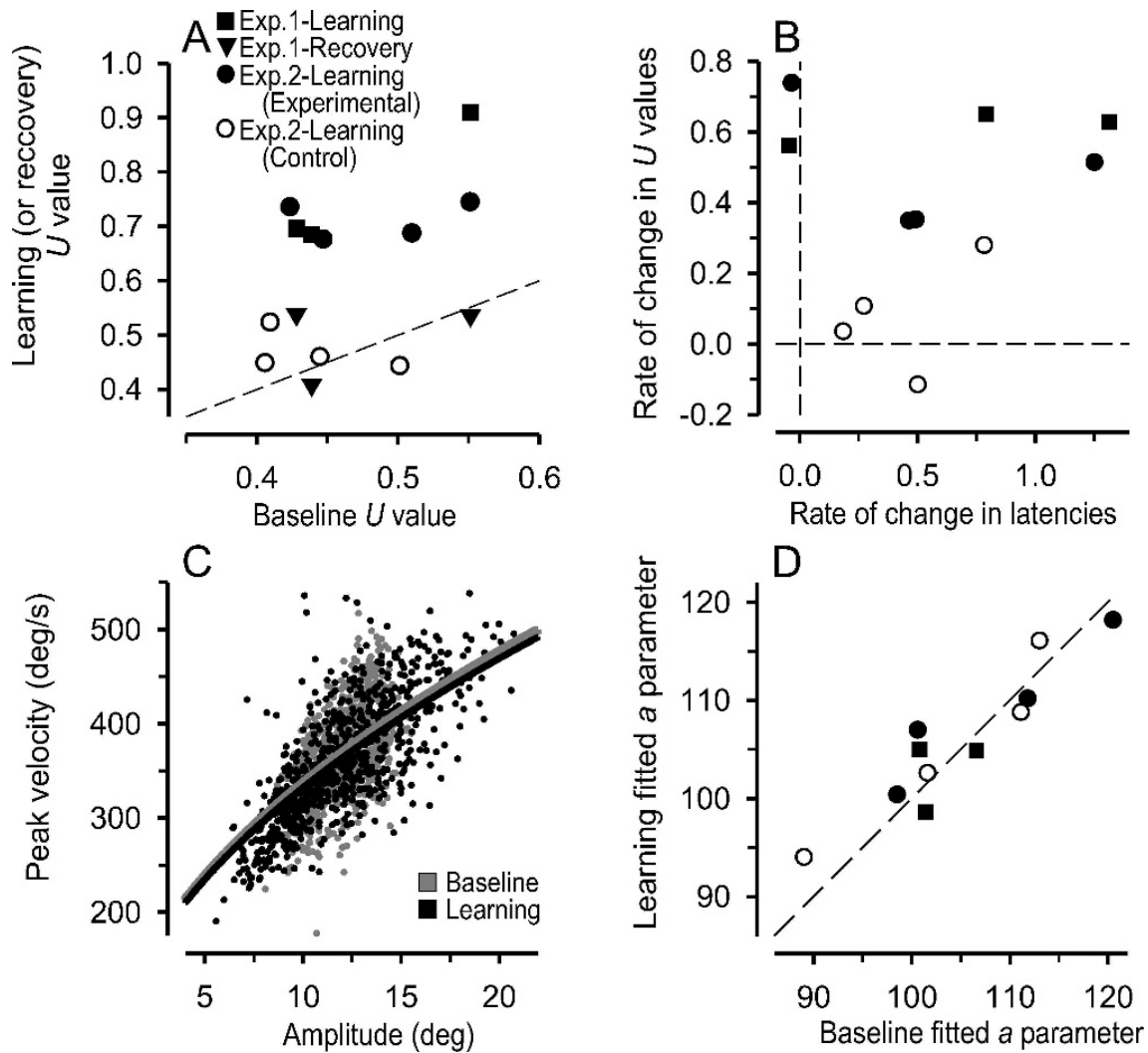


Fig. 4. A: U values during baseline with respect to U values during learning (circles and squares) or recovery (down triangles). Circles, data from Experiment 2. Filled circles, experimental participants. Unfilled circles, control participants. B: Rates of change (end of learning with respect to end of stabilized baseline) in U values as a function of rates of change in latencies. Symbols as in panel A. C: Example of the relationships between saccadic amplitude and peak velocity (participant S2). Gray dots, baseline regular saccades. Black dots, stabilized saccades of the last four learning sessions. Solid lines indicate the corresponding fits. D: Main sequence fitted a parameters at the end of learning as a function of the a parameters during regular baseline. Symbols as in panel A.

increased and then reduced dispersions of saccadic and manual latencies distributions: In a discrimination task, correct responses (saccades or button presses) were reinforced depending on latency variability criteria. Together with our present study, these experiments suggest that variability levels of various saccadic dimensions can be independently controlled by reinforcement contingencies. An important feature of these results is that they do not support the classical view stating that sensorimotor variability originates exclusively from some uncontrollable stochastic internal noise (Faisal et al., 2008; Van Beers, 2007, 2008).

One could argue that the increased oculomotor variability might be achieved through specific strategies—that is, (conscious or not) cognitive processes involved in visual processing, target selection and motor programming depending on stimulus properties and stored information (Glimcher, 2003)—that would affect various response parameters. Saccadic latencies are often viewed as reflecting such decision-making strategies (Reddi & Carpenter, 2000). To probe possible effects of our procedure on saccadic reaction time, we computed the rates of change in latencies ([median obtained over the last learning session - median obtained over the last

stabilized baseline session] / median obtained over the last stabilized baseline session). Figure 4B plots these rates of change against those of U values, revealing that latencies increased after learning (from 179 to 277 ms on average over the last sessions, across all participants; Wilcoxon signed-rank test, $p = .005$). Nevertheless, for the experimental participants the correlation coefficient between these measures was -0.057 , revealing that the changes in amplitude variability are not correlated to the changes in latencies ($p = .903$). Moreover, it can be seen that the experimental and control groups differed in the rates of change in their U values but not in those of their latencies (Figure 4B). Even though the possible use of strategies cannot be entirely ruled out, we did not find any consistent changes in latencies.

Saccadic duration, average velocity, and peak velocity all increase as the size of the saccade increases, a relationship known as the “main sequence” (Bahill, Clark & Stark, 1975). Given the results obtained in studies in which saccades were reinforced (Takikawa et al., 2002; Xu-Wilson et al., 2009), one could also expect a change in peak velocities (the maximum eye velocities during saccades) induced by our procedure. We examined the relationships between amplitude and peak velocity: Saccades recorded during the first regular baseline were compared to saccades from the last four learning sessions. Figure 4C depicts an example of these comparisons for participant S2. The relationships were fitted using the following equation (Lebedev, Van Gelder, & Tsui, 1996):

$$a = \frac{\text{Peak Velocity}}{\sqrt{\text{Saccade Amplitude}}}$$

The a parameters before and after learning were similar (106.6 and 104.9, respectively). This lack of change in the a value was true for all participants, as shown Figure 4D (104.9 and 105.9 on average over the first baseline and the last four learning sessions, respectively; Wilcoxon signed-rank test, $p = .401$), indicating that our reinforcement procedures did not alter saccade dynamics even though variability changed.

Optimal control theory of motor control predicts that specific cost functions are minimized by the central nervous system to

generate movements in the most efficient way (Bays & Wolpert, 2007; Todorov, 2004; Wolpert & Ghahramani, 2000). These theoretical propositions have been applied to the saccadic system (Harris, 1995; Harris & Wolpert, 1998, 2006; Van Beers, 2008). Harris (1995) focused particularly on two parameters to account for saccadic control. The first cost is related to vision impairments during saccadic movement: Saccades landing beyond targets must be avoided because they increase the total eye displacement and therefore saccade durations, in which case vision is lost for a period of time longer than necessary. The other cost is due to saccades’ inaccuracy: Gain must be as close to unity as possible to place the retinal image of the target on the high-acuity area of the retina. Harris’ saccadic flight time minimization model predicts that the visual system would reduce the overall saccadic gain given the presence of endpoint variability in order to minimize the proportion of saccades landing beyond the target. There would be a mechanical relation between gain and variability such that when variability is high, the gain must be lower than when the spread of saccadic error is low. In our experiments we found that median gains significantly decreased in 3 experimental participants (S1, S2 and S5), tended to decrease in 2 (S4 and S6) and to increase in 2 others (S3 and S7). By contrast, variability significantly increased in all 7 experimental participants. In the control experiment we observed a decrease in gain in 1 participant (S8) and an increase in the other 3 while variability decreased in 1 participant (S11) and increased in the other 3. Harris’s prediction was observed in 5 participants (S1, S2, S5, S8 and S11) but not in the other 6. In 2 control participants (S9 and S10) the opposite was true: The median and standard deviation both increased. It is, however, noteworthy that when large changes in variability are induced such as in the experimental participants, the median gains did not decrease in 4 out of the 7 cases. In these participants we did not find a correlation between the rates of change in median gains and in U values (obtained for each participant over the last four learning sessions with respect to the stabilized baseline) ($r = -0.322$, $p = 0.481$, NS). This lack of consistent relationship between changes in

gain and variability confirms that we were able to manipulate each parameter independently, a result that challenges Harris' optimization model.

Variability in saccadic amplitude is thought to strongly affect saccadic control in infants. Harris (1995) suggests that during the first months of life cortical structures responsible for the mapping between target eccentricity and saccadic motor output would not be functional. In fact, infants' saccadic gain has been reported to be extremely low (Aslin & Salapatek, 1975). As a consequence, only the cost related to the saccadic duration would be relevant and priority would first be given to minimize saccadic duration. Only after sufficient maturation of these structures could the saccadic system establish the correspondence between visual signals and motor output. The costs associated with saccadic accuracy would then play a role in the saccadic control which would explain the decreased variability and increased gain observed during development. In other words, the change in gain associated with a decreased variability would be driven by cortical maturation. However, our results, which provide further evidence that eye movements have attributes of operant behavior, suggest that the increase in gain and the reduction in variability would be controlled by a single reinforcement process. Saccadic control might be in part learned through the progressive modification of the visual consequences that exert selective pressure on eye movements. Because infants' retinas are far from mature (Yuodelis & Hendrickson, 1986) and visual interactions are limited, the oculomotor system would first tolerate large end-point errors. During the first months of life most eye movements would therefore not be differentially selected. Progressively, clearer and clearer vision might extinguish the saccades that land away from visual objects and select the saccades that lead to a better visual perception. We propose that both the spread and median of saccadic amplitude distributions develop by means of selection processes: Eye movements allowing sufficiently efficient vision to favor visually driven interactions with the environment would be selected, while less efficient movements would be extinguished.

The operant behavior selection theory postulates that variability is necessary for adaptation (Donahoe, Burgos, & Palmer, 1993;

Skinner, 1981). This hypothesis has been probed in experiments on response sequences showing that only after reinforcement contingent on variability can animals emit difficult-to-learn target sequences, unlike subjects in whom variability had not been reinforced (Grunow & Neuringer, 2002; Neuringer, 1993; Neuringer, Deiss, & Olson, 2000; but see Maes & Van der Goot, 2006, and Neuringer, 2009, for a discussion). One may postulate that some variation in saccadic amplitude is necessary to adapt to changes affecting the oculomotor system. For instance, we experimentally reinforced modifications in saccadic gain (Madelain et al., 2008) and suggested that, in real life, saccades that lead to a clear vision of the target are selected by their consequences. This adaptation would require some variability as a basis for differential reinforcement. That environmental contingencies control amplitude variability might thus be critical to maintain saccadic accuracy during the lifespan.

REFERENCES

- Abel, L. A., Schmidt, D., Dell'Osso, L. F., & Daroff, R. B. (1978). Saccadic system plasticity in humans. *Annals of Neurology*, 4, 313–318.
- Abreu-Rodrigues, J., Lattal, K. A., dos Santos, C. V., & Matos, R. A. (2005). Variation, repetition, and choice. *Journal of the Experimental Analysis of Behavior*, 83, 147–168.
- Aslin, R. N., & Salapatek, P. (1975). Saccadic localization of visual targets by the very young human infant. *Perception & Psychophysics*, 17, 293–302.
- Bahcall, D. O., & Kowler, E. (2000). The control of saccadic adaptation: implications for the scanning of natural visual scenes. *Vision Research*, 40, 2779–2796.
- Bahill, A. T., Clark, M. R., & Stark, L. (1975). The main sequence, a tool for studying human eye movements. *Mathematical Biosciences*, 24, 191–204.
- Bays, P. M., & Wolpert, D. M. (2007). Computational principles of sensorimotor control that minimize uncertainty and variability. *The Journal of Physiology*, 578, 387–396.
- Blough, D. S. (1966). The reinforcement of least-frequent interresponse times. *Journal of the Experimental Analysis of Behavior*, 9, 581–591.
- Chen-Harris, H., Joiner, W. M., Ethier, V., Zee, D. S., & Shadmehr, R. (2008). Adaptive control of saccades via internal feedback. *Journal of Neuroscience*, 28, 2804–2813.
- Darcheville, J.-C., Madelain, L., Buquet, C., Charlier, J., & Miossec, Y. (1999). Operant conditioning of the visual smooth pursuit in young infants. *Behavioural Processes*, 46, 131–139.
- Donahoe, J. W., Burgos, J. E., & Palmer, D. C. (1993). A selectionist approach to reinforcement. *Journal of the Experimental Analysis of Behavior*, 60, 17–40.

- Efron, B., Jolivet, E., & Hordan, R. (1995). *Le bootstrap et ses applications: Discrimination & régression*. Saint-Mandé (France): CISIA.
- Eisenberger, R., Armeli, S., & Pretz, J. (1998). Can the promise of reward increase creativity? *Journal of Personality and Social Psychology*, 74, 704–714.
- Faisal, A. A., Selen, L. P., & Wolpert, D. M. (2008). Noise in the nervous system. *Nature Reviews Neuroscience*, 9, 292–303.
- Galbicka, G. (1994). Shaping in the 21st century: Moving percentile schedules into applied settings. *Journal of Applied Behavior Analysis*, 27, 739–760.
- Galbicka, G., Kautz, M. A., & Jagers, T. (1993). Response acquisition under targeted percentile schedules: A continuing quandary for molar models of operant behavior. *Journal of the Experimental Analysis of Behavior*, 60, 171–184.
- Glimcher, P. W. (2003). The neurobiology of visual-saccadic decision making. *Annual Review of Neuroscience*, 26, 133–179.
- Goetz, E. M., & Baer, D. M. (1973). Social control of form diversity and the emergence of new forms in children's blockbuilding. *Journal of Applied Behavior Analysis*, 6, 209–217.
- Grunow, A., & Neuringer, A. (2002). Learning to vary and varying to learn. *Psychonomic Bulletin & Review*, 9, 250–258.
- Harris, C. M. (1995). Does saccadic undershoot minimize saccadic flight-time? A Monte-Carlo study. *Vision Research*, 35, 691–701.
- Harris, C. M., & Wolpert, D. M. (1998). Signal-dependent noise determines motor planning. *Nature*, 394, 780–784.
- Harris, C. M., & Wolpert, D. M. (2006). The main sequence of saccades optimizes speed-accuracy trade-off. *Biological Cybernetics*, 95, 21–29.
- Havermann, K., & Lappe, M. (2010). The influence of the consistency of post-saccadic visual error on saccadic adaptation. *Journal of Neurophysiology*, 103, 3302–3310.
- Henson, D. B. (1979). Investigation into corrective saccadic eye movements for refixation amplitudes of 10 degrees and below. *Vision Research*, 19, 57–61.
- Hopkinson, J., & Neuringer, A. (2003). Modifying behavioral variability in moderately depressed students. *Behavior Modification*, 27, 251–264.
- Hopp, J. J., & Fuchs, A. F. (2004). The characteristics and neuronal substrate of saccadic eye movement plasticity. *Progress in Neurobiology*, 72, 27–53.
- Ikeda, T., & Hikosaka, O. (2003). Reward-dependent gain and bias of visual responses in primate superior colliculus. *Neuron*, 39, 693–700.
- Ingvaldsen, R. P., & Whiting, H. T. (1997). Modern views on motor skill learning are not 'representative'! *Human Movement Science*, 16, 705–732.
- Joyce, J. H., & Chase, P. N. (1990). Effects of response variability on the sensitivity of rule-governed behavior. *Journal of the Experimental Analysis of Behavior*, 54, 251–262.
- Kommerell, G., Olivier, D., & Theopold, H. (1976). Adaptive programming of phasic and tonic components in saccadic eye movements. Investigations of patients with abducens palsy. *Investigative Ophthalmology*, 15, 657–660.
- Lauwereyns, J., Watanabe, K., Coe, B., & Hikosaka, O. (2002). A neural correlate of response bias in monkey caudate nucleus. *Nature*, 418, 413–417.
- Lebedev, S., Van Gelder, P., & Tsui, W. H. (1996). Square-root relations between main saccadic parameters. *Investigative Ophthalmology & Visual Science*, 37, 2750–2758.
- Lee, R., Sturmey, P., & Fields, L. (2007). Schedule-induced and operant mechanisms that influence response variability: a review and implications for future investigations. *The Psychological Record*, 57, 429–455.
- Machado, A. (1989). Operant conditioning of behavioral variability using a percentile reinforcement schedule. *Journal of the Experimental Analysis of Behavior*, 52, 155–166.
- Machado, A. (1992). Behavioral variability and frequency-dependent selection. *Journal of the Experimental Analysis of Behavior*, 58, 241–263.
- Machado, A. (1997). Increasing the variability of response sequences in pigeons by adjusting the frequency of switching between two keys. *Journal of the Experimental Analysis of Behavior*, 68, 1–25.
- Madelain, L., Champrenaut, L., & Chauvin, A. (2007). Control of sensorimotor variability by consequences. *Journal of Neurophysiology*, 98, 2255–2265.
- Madelain, L., & Krauzlis, R. J. (2003). Effects of learning on smooth pursuit during transient disappearance of a visual target. *Journal of Neurophysiology*, 90, 972–982.
- Madelain, L., Paeye, C., & Wallman, J. (2008). Saccadic adaptation: reinforcement can drive motor adaptation. *Journal of Vision*, 8(6), 919.
- Maes, J. H. R., & Van der Goot, M. (2006). Human operant learning under concurrent reinforcement of response variability. *Learning and Motivation*, 37, 79–92.
- Martinez, W. L., & Martinez, A. R. (2002). *Computational Statistics Handbook with Matlab*. Boca Raton, FL: Chapman and Hall.
- McLaughlin, S. (1967). Parametric adjustment in saccadic eye movements. *Perception & Psychophysics*, 2, 359–362.
- Miller, N., & Neuringer, A. (2000). Reinforcing variability in adolescents with autism. *Journal of Applied Behavior Analysis*, 33, 151–165.
- Montagnini, A., & Chelazzi, L. (2005). The urgency to look: prompt saccades to the benefit of perception. *Vision Research*, 45, 3391–3401.
- Nakamura, K., & Hikosaka, O. (2006). Role of dopamine in the primate caudate nucleus in reward modulation of saccades. *Journal of Neuroscience*, 26, 5360–5369.
- Neuringer, A. (1986). Can people behave "randomly?": the role of feedback. *Journal of Experimental Psychology: General*, 115, 62–75.
- Neuringer, A. (1992). Choosing to vary and repeat. *Psychological Science*, 3, 246–250.
- Neuringer, A. (1993). Reinforced variation and selection. *Animal Learning & Behavior*, 21(2), 83–91.
- Neuringer, A. (2002). Operant variability: evidence, functions, and theory. *Psychonomic Bulletin & Review*, 9, 672–705.
- Neuringer, A. (2009). Operant variability and the power of reinforcement. *The Behavior Analyst Today*, 10(2), 319–343.
- Neuringer, A., Deiss, C., & Olson, G. (2000). Reinforced variability and operant learning. *Journal of Experimental Psychology: Animal Behavior Processes*, 26, 98–111.
- Noto, C. T., & Robinson, F. R. (2001). Visual error is the stimulus for saccade gain adaptation. *Cognitive Brain Research*, 12, 301–305.

- Optican, L. M., Zee, D. S., & Chu, F. C. (1985). Adaptive response to ocular muscle weakness in human pursuit and saccadic eye movements. *Journal of Neurophysiology*, 54, 110–122.
- Page, S., & Neuringer, A. (1985). Variability is an operant. *Journal of Experimental Psychology: Animal Behavior Processes*, 11, 429–452.
- Pelisson, D., Alahyane, N., Panouilleres, M., & Tilikete, C. (2010). Sensorimotor adaptation of saccadic eye movements. *Neuroscience and Biobehavioral Reviews*, 34, 1103–1120.
- Pryor, K. W., Haag, R., & O'Reilly, J. (1969). The creative porpoise: training for novel behavior. *Journal of the Experimental Analysis of Behavior*, 12, 653–661.
- Reddi, B. A., & Carpenter, R. H. (2000). The influence of urgency on decision time. *Nature Neuroscience*, 3, 827–830.
- Robinson, F. R., Noto, C. T., & Bevans, S. E. (2003). Effect of visual error size on saccade adaptation in monkey. *Journal of Neurophysiology*, 90, 1235–1244.
- Robinson, F. R., Soetedjo, R., & Noto, C. (2006). Distinct short-term and long-term adaptation to reduce saccade size in monkey. *Journal of Neurophysiology*, 96, 1030–1041.
- Sato, M., & Hikosaka, O. (2002). Role of primate substantia nigra pars reticulata in reward-oriented saccadic eye movement. *Journal of Neuroscience*, 22, 2363–2373.
- Schultz, W. (2000). Multiple reward signals in the brain. *Nature Neuroscience*, 1, 199–207.
- Seeberger, T., Noto, C., & Robinson, F. (2002). Non-visual information does not drive saccade gain adaptation in monkeys. *Brain Research*, 956, 374–379.
- Skinner, B. F. (1981). Selection by consequences. *Science*, 213, 501–504.
- Stokes, P. D. (1995). Learned variability. *Animal Learning & Behavior*, 23, 164–176.
- Stokes, P. D., & Balsam, P. (2001). An optimal period for setting sustained variability levels. *Psychonomic Bulletin & Review*, 8, 177–184.
- Takikawa, Y., Kawagoe, R., Itoh, H., Nakahara, H., & Hikosaka, O. (2002). Modulation of saccadic eye movements by predicted reward outcome. *Experimental Brain Research*, 142, 284–291.
- Todorov, E. (2004). Optimality principles in sensorimotor control (review). *Nature Neuroscience*, 7, 907–915.
- Van Beers, R. J. (2007). The sources of variability in saccadic eye movements. *Journal of Neuroscience*, 27, 8757–8770.
- Van Beers, R. J. (2008). Saccadic Eye Movements Minimize the Consequences of Motor Noise. *PLoS ONE* 3, e2070, Retrieved August 17, 2010, from <http://www.plosone.org/article/>
- Wallman, J., & Fuchs, A. F. (1998). Saccadic gain modification: visual error drives motor adaptation. *Journal of Neurophysiology*, 80, 2405–2416.
- Watanabe, K., Lauwereyns, J., & Hikosaka, O. (2003). Neural correlates of rewarded and unrewarded eye movements in the primate caudate nucleus. *Journal of Neuroscience*, 23, 10052–10057.
- Wolpert, D. M., & Ghahramani, Z. (2000). Computational principles of movement neuroscience. *Nature Neuroscience*, 3, 1212–1217.
- Xu-Wilson, M., Zee, D. S., & Shadmehr, R. (2009). The intrinsic value of visual information affects saccade velocities. *Experimental Brain Research*, 196, 475–481.
- Yuodelis, C., & Hendrickson, A. (1986). A qualitative and quantitative analysis of the human fovea during development. *Vision Research*, 26, 847–855.
- Zee, D. S., Optican, L. M., Cook, J. D., Robinson, D. A., & Engel, W. K. (1976). Slow saccades in spinocerebellar degeneration. *Archives of Neurology*, 33, 243–251.

Received: August 17, 2010

Final Acceptance: November 1, 2010

